

STIC-Biotech/ChemLib

178033

From: Whiteman, Brian
Sent: Monday, January 30, 2006 3:14 PM
To: STIC-Biotech/ChemLib
Subject: seq search

10/069386

SEQ ID NOs: 1 and 2

1) interference search

2) us patent and published us patent application databases

Thank you,

Brian Whiteman
Remsen, 2D14
mail box 2C18
Patent Examiner - Art Unit 1635
United States Patent and Trademark Office
(571) 272-0764

RECEIVED
JAN 30 2006
STIC

Searcher: _____
Searcher Phone: _____
Date Searcher Picked up: _____
Date completed: _____
Searcher Prep Time: _____
Online Time: _____

Type of Search
NA# _____ AA# _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure #: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable
STN: _____
DIALOG: _____
QUESTEL/ORBIS: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other (Specify): _____

This Page Blank (uspio)

ORIGIN	
Query Match	100.0%; Score 591; DB 8; Length 1542;
Best Local Similarity	100.0%; Pred. No. 8.2e-114;

PI TOSHIO OTA, TETSUO NISHIKAWA, TAKAO ISOGAI, KOJI HAYASHI, SHIZUO
PI ISHII,
PI YORI KANAI, AI WAKAMATSU, TOMOYASU SUGIYAMA, KEIICHI NAGAI, PI
SHINICHI KOJIMA,
PI TETSUO OTSUKI, HISASHI KOGA
PC
C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC

10,
PC C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key

Location/Qualifiers
(93) .. (680).

FEATURES
source Location/Qualifiers
1..1325
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 99.7%; Score 589.4; DB 6; Length 1325;
Best Local Similarity 99.8%; Pred. No. 1.8e-113;
Matches 590; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 ATGAGGGAGGCTTGAAGGAAACACTGTGATTGGAAGAGAGAGAGAGTGGAG 60
|||
93 ATGTGGAGGCTTGAAGGAAACACTGTGATTGGAAGAGAGAGAGAGTGGAG 152
61 TGAAGTCCAGAGGCTTCAAGACTACAGAAAGCCCTGCTCCGATCTCCCTAAGCAA 120
|||
153 TGAAGTCCAGAGGCTTCAAGACTACAGAAAGCCCTGCTCCGATCTCCCTAAGCAA 212
121 GTCCAGCGAGCTTGGGCCCCCAGAGACCCAGACTCCGAGGCAATGCTCATTCATAAC 180
213 GTCCAGCGAGCTTGGGCCCCCAGAGACCCAGACTCCGAGGCAATGCTCATTCATAAC 272
181 ACCCTTCAACAGCTGAGGCTGCACTTGGCTGGCTCCGCTCCGCTCCGCTCCGCTCCG 240
273 ACCCTTCAACAGCTGAGGCTGCACTTGGCTGGCTCCGCTCCGCTCCGCTCCGCTCCG 332
241 CCCCTTCTCTGGGGGAGAGGATTTCTCCCTGTAGAGCCACATTTGGCTCTATCTCCAG 300
333 CCCCTTCTCTGGGGGAGAGGATTTCTCCCTGTAGAGCCACATTTGGCTCTATCTCCAG 392
301 GAGCTGACACCTTCATGATGAGGACTAGAGCCCTCAGAAATCCAGTACTCCCTTGGC 360
393 GAGCTGACACCTTCATGATGAGGACTAGAGCCCTCAGAAATCCAGTACTCCCTTGGC 452
361 CTCCAGAAATGAAGTGCACCCCAAGCTTGATCAGTCTTTTGAAGAGCTGTAGCTCCGG 420
453 CTCCAGAAATGAAGTGCACCCCAAGCTTGATCAGTCTTTTGAAGAGCTGTAGCTCCGG 512
421 TACTTGGGGGAGCTTGGGCTGTAGTACTTTCTTGGACATGAGACATCTGGGGTAGAA 480
513 TACTTGGGGGAGCTTGGGCTGTAGTACTTTCTTGGACATGAGACATCTGGGGTAGAA 572
481 AAGAGCTGTGACGGGCCCCCAGAGGCTCTCAACAACCTTCTGTGCCCCAGATTCT 540
573 AAGAGCTGTGACGGGCCCCCAGAGGCTCTCTCAACAACCTTCTGTGCCCCAGATTCT 632
541 TGGAGTGAATGAATGATGATCAATCATGAAATCATTTCTGGGGTCTCTAA 591
633 TGGAGTGAATGAATGATGATCAATCATGAAATCATTTCTGGGGTCTCTAA 683

RESULT 4
CQ783944 1325 bp DNA linear PAT 17-MAR-2004
LOCUS CQ783944
DEFINITION Sequence 4084 from Patent EP1396543.
ACCESSION CQ783944
VERSION CQ783944.1 GI:45503845
KEYWORDS
ORGANISM
Homo sapiens (human)
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.

REFERENCE
AUTHORS
Ota, T., Nishikawa, T., Isogai, T., Hayashi, K., Ishii, S., Kawai, Y.,
Makamatsu, A., Sugiyama, T., Nagai, K., Kojima, S., Otsuki, T. and
Koga, H.

TITLE
JOURNAL
Primer for synthesizing full length cDNA clones and their use
Patent: EP 1396543-A 4084 10-MAR-2004;
Research Association for Biotechnology (JP)

FEATURES
source Location/Qualifiers
1..1325
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
93..683
/note="unnamed protein product"

CDS

/protein_id="CAP6937.1"
/db_xref="GI:45503846"
/translation="MWGILKRHSIDLEBEREWEWSPAGLQSYQALIRISIDKYORS
LGRAPSLRHYLIINTIQQALALAPALPEPLPGEDESLATIGSILREL
DTSMGTBPQNPVPTPLGIONEVPPOPPOPEPLFALSRLVDSGLDFFLDITSAVE
KEPAPAPPEPPIINLFCAGSNEMNLDHIMEITLLS"

ORIGIN

Query Match 99.7%; Score 589.4; DB 6; Length 1325;
Best Local Similarity 99.8%; Pred. No. 1.8e-113;
Matches 590; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 ATGAGGGAGGCTTGAAGGAAACACTGTGATTGGAAGAGAGAGAGAGTGGAG 60
|||
93 ATGTGGAGGCTTGAAGGAAACACTGTGATTGGAAGAGAGAGAGAGTGGAG 152
61 TGAAGTCCAGAGGCTTCAAGACTACAGAAAGCCCTGCTCCGATCTCCCTAAGCAA 120
|||
153 TGAAGTCCAGAGGCTTCAAGACTACAGAAAGCCCTGCTCCGATCTCCCTAAGCAA 212
121 GTCCAGCGAGCTTGGGCCCCCAGAGACCCAGACTCCGAGGCAATGCTCATTCATAAC 180
213 GTCCAGCGAGCTTGGGCCCCCAGAGACCCAGACTCCGAGGCAATGCTCATTCATAAC 272
181 ACCCTTCAACAGCTGAGGCTGCACTTGGCTGGCTCCGCTCCGCTCCGCTCCGCTCCG 240
273 ACCCTTCAACAGCTGAGGCTGCACTTGGCTGGCTCCGCTCCGCTCCGCTCCGCTCCG 332
241 CCCCTTCTCTGGGGGAGAGGATTTCTCCCTGTAGAGCCACATTTGGCTCTATCTCCAG 300
333 CCCCTTCTCTGGGGGAGAGGATTTCTCCCTGTAGAGCCACATTTGGCTCTATCTCCAG 392
301 GAGCTGACACCTTCATGATGAGGACTAGAGCCCTCAGAAATCCAGTACTCCCTTGGC 360
393 GAGCTGACACCTTCATGATGAGGACTAGAGCCCTCAGAAATCCAGTACTCCCTTGGC 452
361 CTCCAGAAATGAAGTGCACCCCAAGCTTGATCAGTCTTTTGAAGAGCTGTAGCTCCGG 420
453 CTCCAGAAATGAAGTGCACCCCAAGCTTGATCAGTCTTTTGAAGAGCTGTAGCTCCGG 512
421 TACTTGGGGGAGCTTGGGCTGTAGTACTTTCTTGGACATGAGACATCTGGGGTAGAA 480
513 TACTTGGGGGAGCTTGGGCTGTAGTACTTTCTTGGACATGAGACATCTGGGGTAGAA 572
481 AAGAGCTGTGACGGGCCCCCAGAGGCTCTCAACAACCTTCTGTGCCCCAGATTCT 540
573 AAGAGCTGTGACGGGCCCCCAGAGGCTCTCTCAACAACCTTCTGTGCCCCAGATTCT 632
541 TGGAGTGAATGAATGATGATCAATCATGAAATCATTTCTGGGGTCTCTAA 591
633 TGGAGTGAATGAATGATGATCAATCATGAAATCATTTCTGGGGTCTCTAA 683

RESULT 5
AK074604 1325 bp mRNA linear PRI 03-SEP-2002
LOCUS AK074604
DEFINITION Homo sapiens cDNA FLJ90123 f1s, clone HEMBA1007226, highly similar
to Homo sapiens RPA-binding trans-activator (RBT1) mRNA.
ACCESSION AK074604
VERSION AK074604.1 GI:22760150
KEYWORDS
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.

REFERENCE 1
Isogai, T., Ota, T., Nishikawa, T., Hayashi, K., Otsuki, T.,
Sugiyama, T., Suzuki, Y., Nagai, K., Sugano, S., Ishii, S.,
Kawai-Hiro, Y., Saito, K., Yamamoto, J., Wakamatsu, A., Nakamura, Y.,
Kojima, S., Nagahara, K., Masuho, Y., Ota, T., Okano, K., Yoshikawa, Y.,
Aotake, S., Sasaki, N., Hattori, A., Okumura, K., Iwayanagi, T. and
Ninomiya, K.
NEO human cDNA sequencing project
Unpublished
2 (bases 1 to 1325)
Isogai, T. and Otsuki, T.
Direct Submission
Submitted (25-MAR-2002) Takao Isogai, Helix Research Institute,
Genomics Laboratory, 1532-3 Yana, Kizazuru, Chiba 292-0812, Japan
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
NEO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology; cDNA library construction:
Institute of Medical Science, University of Tokyo, Laboratory of
Genome Structure, Human Genome Center; cDNA 5'-E 3'-end one pass
sequencing and clone selection: Helix Research Institute (supported
by Japan Key Technology Center etc.).
Location/Qualifiers
1. 1325
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HEMBA1007226"
/tissue_type="whole embryo, mainly head"
/clone_lib="HEMBA1"
/dev_stage="embryo, 10 weeks"
/note="Cloning vector: pMB18SFL3"

FEATURES
source

ORIGIN
Query Match 99.7%; Score 589.4; DB 8; Length 1325;
Best Local Similarity 99.8%; Pred. No. 1.8e-113;
Matches 590; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 ATGAGGAGAGGCTTGAAGAGAAACCTCTGATTTGAAAGAGAGAGAGAGGTGGAG 60
DB 93 ATGAGGAGAGGCTTGAAGAGAAACCTCTGATTTGAAAGAGAGAGAGAGGTGGAG 152
QY 61 TGGAGTCCAGCAGGCTTCAAGAGCTACCGAGAAAGCTCTCCGATCTCCCTAGACAA 120
DB 153 TGGAGTCCAGCAGGCTTCAAGAGCTACCGAGAAAGCTCTCCGATCTCCCTAGACAA 212
QY 121 GTCCAGCAGCAGGCTTCAAGAGCTACCGAGAAAGCTCTCCGATCTCCCTAGACAA 180
DB 213 GTCCAGCAGCAGGCTTCAAGAGCTACCGAGAAAGCTCTCCGATCTCCCTAGACAA 272
QY 181 ACCCTCCAGCAGGCTTCAAGAGCTACCGAGAAAGCTCTCCGATCTCCCTAGACAA 240
DB 273 ACCCTCCAGCAGGCTTCAAGAGCTACCGAGAAAGCTCTCCGATCTCCCTAGACAA 332
QY 241 CCCCTCTTCTGGGAGAGAGATTTCTCCCTGACGACCAATTTGGCTTATTCCTAGAG 300
DB 333 CCCCTCTTCTGGGAGAGAGATTTCTCCCTGACGACCAATTTGGCTTATTCCTAGAG 392
QY 301 GAGCTGACACCTCCATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
DB 333 GAGCTGACACCTCCATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 452
QY 361 CTCCAGAGATGAG 420
DB 453 CTCCAGAGATGAG 512
QY 421 TACTTGGGGAG 480
DB 513 TACTTGGGGAG 572

QY 481 AAGAGAGCTGACAGGAGCCCAACAGAGAGCTCTCTACAACTCTTGTGGCCCAAGTCT 540
DB 573 AAGAGAGCTGACAGGAGCCCAACAGAGAGCTCTCTACAACTCTTGTGGCCCAAGTCT 632

QY 541 TGGGAGTGAATGAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 591
DB 633 TGGGAGTGAATGAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 683

RESULT 6
BC014061
LOCUS 1346 bp mRNA linear PRI 25-JUL-2005
DEFINITION Homo sapiens SERTA domain containing 3, transcript variant 2, mRNA
(cDNA clone MGC:119994 IMAGE:3926937), complete cds.
ACCESSION BC014061
VERSION BC014061.2 GI:33873834
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 1346)
Strausberg, R.L., Feingold, R.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Sherman, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Scheeler, C.F., Blac, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Dietzenko, J.L., Marziani, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Udell, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abrahamson, R.D., Mullahy, S.J., Bosak, S.A., McSwan, P.J.,
McKernan, R.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hult, K.S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Pahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywicki, M.I., Staleka, U., Small, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Mammalian Gene Collection Program Team
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 1346)

CONSTRM
TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS
CONSTRM
TITLE
JOURNAL
REMARK
COMMENT

NIH MGC Project
Direct Submission
Submitted (10-SEP-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgs.nhl.nih.gov>
On Aug 19, 2003 this sequence version replaced gi:15559388.
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DRP
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nih.gov>
Contact: nisc.mgc@nih.gov
Akher, N., Ayala, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Leric, P., Legaspi, R.,
Maduro, O.L., Masello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Stancirjop, S., Thomas, P.J., Touchman, J.W.,
Taurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNLT at: <http://image.llnl.gov>
 Series: IRAL Plate: 28 Row: f Column: 10
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 15718683.

FEATURES

SOURCE

1..1346
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:19994 IMAGE:3926937"
 /tissue_type="Ovary, adenocarcinoma"
 /clone_id="N1H MGC_9"
 /lab_host="DH10B-R"
 /note="Vector: pOTB7"
 1..1346
 /gene="SERTAD3"
 /note="synonym: RBT1"
 /db_xref="GeneID:29946"
 92..682
 /gene="SERTAD3"
 /codon_start=1
 /product="RPA-binding trans-activator"
 /protein_id="AAH14061.1"
 /db_xref="GI:15559389"
 /db_xref="GeneID:29946"
 /translation="MVGGLKRRKSDLEEREREMWSPAGLOSVOQALRLISLQVRS
 LGRAPRLRRVLTHTTQLOALRLAPALPPEPLFLEBEDSLSTIGLIREL
 DTSMDTEPQNPVPLGLONEVPPQPPLEALSRILGSDGLDFLDITSAYE
 KEPARAPPEPHNLFCAPSWEMWELDHIMEILLDS"

CDS

1..1346
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:19994 IMAGE:3926937"
 /tissue_type="Ovary, adenocarcinoma"
 /clone_id="N1H MGC_9"
 /lab_host="DH10B-R"
 /note="Vector: pOTB7"
 1..1346
 /gene="SERTAD3"
 /note="synonym: RBT1"
 /db_xref="GeneID:29946"
 92..682
 /gene="SERTAD3"
 /codon_start=1
 /product="RPA-binding trans-activator"
 /protein_id="AAH14061.1"
 /db_xref="GI:15559389"
 /db_xref="GeneID:29946"
 /translation="MVGGLKRRKSDLEEREREMWSPAGLOSVOQALRLISLQVRS
 LGRAPRLRRVLTHTTQLOALRLAPALPPEPLFLEBEDSLSTIGLIREL
 DTSMDTEPQNPVPLGLONEVPPQPPLEALSRILGSDGLDFLDITSAYE
 KEPARAPPEPHNLFCAPSWEMWELDHIMEILLDS"

ORIGIN

Query Match 99.7%; Score 589.4; DB 8; Length 1346;
 Best Local Similarity 99.8%; Pred. No. 1.8e-113;
 Matches 590; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 ATGGAGGAGGCTTAAGAGGAAACACTGTATTTGGAAGAGAGAGAGAGTGGAG 60
 92 ATGGAGGAGGCTTAAGAGGAAACACTGTATTTGGAAGAGAGAGAGAGTGGAG 151
 61 TGGAGTCCAGCAGGCTTCAAGAGTACAGCAGCAGGCTTCCGATCTCCCTAGACAA 120
 152 TGGAGTCCAGCAGGCTTCAAGAGTACAGCAGCAGGCTTCCGATCTCCCTAGACAA 211
 121 GTCCAGGCGAGCTTGGGCCCCGAGACCCAGGCTTCCGAGGAGTCTCTATCTAAC 180
 212 GTCCAGGCGAGCTTGGGCCCCGAGACCCAGGCTTCCGAGGAGTCTCTATCTAAC 271
 181 ACCCTCCAAAGCTGACAGGCTGACCTTGGGCTCCGAGCCCGCCCTGCCCCGAG 240
 212 ACCCTCCAAAGCTGACAGGCTGACCTTGGGCTCCGAGCCCGCCCTGCCCCGAG 331
 241 CCCCTCTTCTGGGCGAGAGATTTCTCCCTGTAGCAGCAGCATTGGCTTATCTCAG 300
 332 CCCCTCTTCTGGGCGAGAGATTTCTCCCTGTAGCAGCAGCATTGGCTTATCTCAG 391
 301 GAGCTGACACTCTCATGTAGTGGACTGAGCCCCCTCAGATTCAGTGACTCCCTTGGC 360
 392 GAGCTGACACTCTCATGTAGTGGACTGAGCCCCCTCAGATTCAGTGACTCCCTTGGC 451
 361 CTCGAGATGAGAGTCAACCCGAGCTGATTCAGTCTTCTTGAAGCTCTGAGCTCCGG 420
 452 CTCGAGATGAGAGTCAACCCGAGCTGATTCAGTCTTCTTGAAGCTCTGAGCTCCGG 511
 421 TACTTGGGGAATCTTGGCTGTAGTACTTCTTCTTGAACATTGACACATCTGGGTAGA 480
 512 TACTTGGGGAATCTTGGCTGTAGTACTTCTTCTTGAACATTGACACATCTGGGTAGA 571
 481 AAGAGGCTGACGCGGCCCCACACAGGCTCTCTACACACTTCTTGTGCCCCAGTTCT 540
 572 AAGAGGCTGACGCGGCCCCACACAGGCTCTCTCTCAACACTTCTGTGCCCCAGTTCT 631

541 TGGAGTGAATGAACCTGATCAGTCACTGGAATATCATTTCTGGGCTCTTA 591
 632 TGGAGTGAATGAACCTGATCAGTCACTGGAATATCATTTCTGGGCTCTTA 682

RESULT 7

LOCUS

BC050643 1420 bp mRNA linear PRI 08-MAR-2005
 Homo sapiens SERRA domain containing 3, transcript variant 2, mRNA
 (CDNA clone MGC:60134 IMAGE:6503515), complete cds.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

1 (bases 1 to 1420)
 Strausberg RL, Feingold EA, Grouse LH, Derge JG, Klausner RD,
 Collins FS, Wagner L, Shenmen CM, Schuler GD, Altschul SF, Zeeberg
 B, Buetow KH, Schaefer CF, Bhat NK, Hopkins RF, Jordan H, Moore T,
 Max SI, Wang J, Haich P, Diatchenko L, Marusina K, Farmer AA, Rubin
 GM, Hong L, Stapleton M, Soares MB, Bonaldo MP, Casavant TL,
 Scheetz TE, Brownstein MJ, Ucdin TB, Toshiyuki S, Carninci P,
 Prange C, Raha S, Loquellano NA, Peters GJ, Abramson RD, Mullahy
 SJ, Bock SA, McEwan PJ, McKernan PJ, Malek JA, Gunaratne PH,
 Richards S, Worley KC, Hale S, Garcia AM, Gay LJ, Hulyk SM,
 Villalón DK, Muzny DM, Sodergren EJ, Lu X, Gibbs RA, Bahay J,
 Helton E, Kettelman M, Madan A, Rodriguez S, Sanchez A, Whitting M,
 Madan A, Young AC, Shevchenko Y, Boulford GJ, Blakesley RW,
 Touchman JW, Green ED, Dickinson MC, Rodriguez AC, Grimwood J,
 Schmutz J, Myers RM, Butlerfield YS, Rodriguez MI, Skalka U,
 Smalusz DE, Scherch A, Schein JE, Jones SJ and Marra MA.
 Mammalian Gene Collection Program Team
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 2 (bases 1 to 1420)
 Director MGC Project.
 Direct Submission
 Submitted (08-APR-2003) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA

REMARK

COMMENT

NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
 On Aug 25, 2003 this sequence version replaced gi:29791864.
 Contact: MGC help desk
 Email: cgapbe-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNLT)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www-shgc.stanford.edu>
 Contact: (Dickson Mark) mcdpaxll.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
 R. M.

FEATURES

SOURCE

1..1420
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:60134 IMAGE:6503515"
 /tissue_type="Ovary, adenocarcinoma"
 /clone_id="N1H MGC_71"

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNLT at: <http://image.llnl.gov>
 Series: IRAL Plate: 110 Row: o Column: 24
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 15718683.

```

gene
1. .1420
/gene="SERTAD3"
/note="synonym: RBT1"
/db_xref="Geneid:29946"
168. .758
CDS
/lab_host="DH10B"
/note="vector: pCMV-SPORT6"

```

```

/product="RPA-binding Trans-activator"
/protein_id="AAH50643.1"
/db_xref="GI:29791865"
/db_xref="GeneID:29946"
/translation="MVGGLKRRKSDIEEERKRWSPAGLOSYYQALRLISLQVQRK
LGGPAAKRRHVLHLLTQQLAKLIALPAPLPPPLPLGGSDLSLSTATGSLIKRSI
DESDAPPEPQNPVGLQGNLNPVGLQGNLNPVGLQGNLNPVGLQGNLNPVGLQGNL
LEPPAAKPEEPHNLCAKPGSWENNELDPIHMLITLSS"

```

ORIGIN

Query Match	99.7%	Score 589.4;	DB 8;	Length 1420;
Best Local Similarity	99.8%	Pred. No. 1.8e-113;		
Matches 590;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;

QY	1	ATGAGGAGGCTTGAAAGAGAAACACTGATTTTGAAAGAGAGAGAGAGGTGGAG	60
Db	168	ATGATGGAGGCTTGAAGAGAAACACTGATTTTGAAAGAGAGAGAGAGGTGGAG	227
QY	61	TGAGTCCAGCAGGCTTCAAGCTACAGCAGACGCTGCTCCGATCTCCTTAGCAAA	120
Db	228	TGAGTCCAGCAGGCTTCAAGCTACAGCAGACGCTGCTCCGATCTCCTTAGCAAA	287
QY	121	GTCCAGGCAACCTGGGGCCCCCGAGCAACCAAGCTCCGAGGCAATGTCATCATCAAC	180
Db	288	GTCCAGGCAACCTGGGGCCCCCGAGCAACCAAGCTCCGAGGCAATGTCATCATCAAC	347
QY	181	ACCTCCAAACAGCTGCAGAGCTGCACTTGAGCTGCCGCCCTGCCTGCCCCCAG	240
Db	348	ACCTCCAAACAGCTGCAGAGCTGCACTTGAGCTGCCGCCCTGCCTGCCCCCAG	407
QY	241	CCCTCTTCTGGGGGAGAGATTCTCCCTGCAAGCCACATTGGCTATCTCTCAG	300
Db	408	CCCTCTTCTGGGGGAGAGATTCTCCCTGCAAGCCACATTGGCTATCTCTCAG	467
QY	301	GAGCTGCACACTCCATGGATGGATGGAGCTGAGCCCCCAGAAATCCAGTGACTCCCTTGGC	360
Db	468	GAGCTGCACACTCCATGGATGGAGCTGAGCCCCCAGAAATCCAGTGACTCCCTTGGC	527
QY	361	CTCCAGAAATGAAGTGCACCCCAAGCCTGATCCAGTCTTTTGAAGCTCTGAGCTCCGG	420
Db	528	CTCCAGAAATGAAGTGCACCCCAAGCCTGATCCAGTCTTTTGAAGCTCTGAGCTCCGG	587
QY	421	TACTTGGGGGACTCTGGCTGTGATGACTTTTCTGCAATTTGACACATCTGCGGTAGAA	480
Db	588	TACTTGGGGGACTCTGGCTGTGATGACTTTTCTGCAATTTGACACATCTGCGGTAGAA	647
QY	481	AAGAGCTTGCACGGGGCCCCACAGAAGCTCTGCACAACTTTCTGTGCCCAAGTTCT	540
Db	648	AAGAGCTTGCACGGGGCCCCACAGAAGCTCTCTCACAACTTTCTGTGCCCAAGTTCT	707
QY	541	TGGGAGTGAATGAATGATGATCAATCAATGAAATCAATTCTGGGGTCTTAA	591
Db	708	TGGGAGTGAATGAATGATGATCAATCAATGAAATCAATTCTGGGGTCTTAA	758

RESULT	8
CQ731553	
LOCUS	CQ731553
DEFINITION	CQ731553
ACCESSION	Sequence 17467 from Patent WO02066579.
VERSION	CQ731553
KEYWORDS	CQ731553.1 GI:42308440
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens

REFERENCE	AUTHORS	TITLE
1	Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.	Kits, such as nucleic acid arrays, comprising a majority of

JOURNAL Patent: WO 02068579-A 17487 06-SEP-2002,
PE Corporation (NY) (US)

FEATURES

Source

```
/organism="Homo sapiens"  
/mol_type="unassigned DNA"  
/db_xref="taxon:9606"
```

ORIGIN

Query Match	99.7%	Score 589.4	DB 6	Length 1536
Best Local Similarity	99.8%	Pred. No. 1,8e-113		
Matches 590	Conservative 0	Mismatches 1	Indels 0	Gaps 0

QY	1	ATGGAGGGAGGCTTGAAGAGGAACACCTCATTTTGAAGAGGAGGAGGAGAGGAG	80
Db	290	ATGGTGGGAGGCTTGAAGAGGAACACTGTATTTGAAGAGGAGGAGGAGGAGGAG	349
QY	61	TGAGATCAGAGGAGGCTTCAAGAGCTACAGCAAGGCTGCTCCGATCTTCCTTAGCAAA	120
Db	350	TGGATGCAGAGGAGGCTTCAAGAGCTACAGCAAGGCTGCTCCGATCTTCCTTAGCAAA	409
QY	121	GTCCAGGCGAGGCTGGGGCCCCCGAGCAACCGAGCTCCGAGGAGCATGTCTCATTCATAAC	180
Db	410	GTCCAGGCGAGGCTGGGGCCCCCGAGCAACCGAGCTCCGAGGAGCATGTCTCATTCATAAC	469
QY	181	ACCCTCCAAAGCTGCAAGGCTGCACTTCGCTGGCTCCGCCCCCTAGCCTGCCCCGAG	240
Db	470	ACCCTCCAAAGCTGCAAGGCTGCACTTCGCTGGCTCCGCCCCCTAGCCTGCCCCGAG	529
QY	241	CCCTCTTCTCTGGGCGAGAGGATTTCTCCCTGTGAGCCACATTGGCTTATCTCGAG	300
Db	530	CCCTCTTCTCTGGGCGAGAGGATTTCTCCCTGTGAGCCACATTGGCTTATCTCGAG	589
QY	301	GAGCTGCAACCTTCANAGGATGGGACGTAGGCCCCCTAGAAATCCAGTGACTCCCTCTGGC	360
Db	590	GAGCTGCAACCTTCANAGGATGGGACGTAGGCCCCCTAGAAATCCAGTGACTCCCTCTGGC	649
QY	361	CTCCAGATGAAGTGCACACCCAGGCTGATCCAGTCTTTGAAGCTCTGAGCTCCCG	420
Db	650	CTCCAGATGAAGTGCACACCCAGGCTGATCCAGTCTTTGAAGCTCTGAGCTCCCG	709
QY	421	TACTTGGGGAGCTCTGGCTGAGTACCTTTCTGCAATTTGACACATCTCGGGTAGAA	480
Db	710	TACTTGGGGAGCTCTGGCTGAGTACCTTTCTGCAATTTGACACATCTCGGGTAGAA	769
QY	481	AAGAGGCTGCAAGGAGCCCCACAGAGGCTCTTCACAACTTTCTGGGCCCAAGTTCT	540
Db	770	AAGAGGCTGCAAGGAGCCCCACAGAGGCTCTTCACAACTTTCTGGGCCCAAGTTCT	829
QY	541	TGGGAGTGAAATGAATCGATCACATCANAGAAATCATTTCTGGGCTCTAA	591
Db	830	TGGGAGTGAAATGAATCGATCACATCANAGAAATCATTTCTGGGCTCTCTAA	880

RESULT 9
AC010271/c AC010271 160643 bp DNA linear PRI 13-JUL-2007
LOCUS Homo sapiens chromosome 19 clone CTC-492K19, complete sequence.
DEFINITION AC010271
ACCESSION AC010271
VERSION AC010271.8 GI:21743753
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

REFERENCE 1 Homidae; Homo.
AUTHORS 1 (bases 1 to 160643)
TITLE DO Joint Genome Institute and Stanford Human Genome Center.
JOURNAL Direct Submission
REFERENCE 2 Unpublished
AUTHORS 2 (bases 1 to 160643)
TITLE DO Joint Genome Institute.
JOURNAL Direct Submission
REFERENCE 3 Submitted (15-SEP-1999) Production Sequencing Facility, DOE Joint
AUTHORS Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
TITLE 3 (bases 1 to 160643)
JOURNAL DO Joint Genome Institute and Stanford Human Genome Center.
REFERENCE 4 Submitted (29-AUG-2000) DOE Joint Genome Institute, 2800 Mitchell
AUTHORS Drive, Walnut Creek, CA 94598, USA
TITLE 4 (bases 1 to 160643)
JOURNAL DO Joint Genome Institute and Stanford Human Genome Center.
REFERENCE 5 Submitted (03-OCT-2001) DOE Joint Genome Institute, 2800 Mitchell
AUTHORS Drive, Walnut Creek, CA 94598, USA
TITLE 5 (bases 1 to 160643)
JOURNAL DO Joint Genome Institute and Stanford Human Genome Center.
REFERENCE 6 Submitted (13-JUL-2002) DOE Joint Genome Institute, 2800 Mitchell
AUTHORS Drive, Walnut Creek, CA 94598, USA
TITLE On Jul 13, 2002 this sequence version replaced gi:15887272.
JOURNAL Draft Sequence Produced by DOE Joint Genome Institute
COMMENT www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.sbgc.stanford.edu
Quality: Phrap Quality >=40 100% of Sequence;
Estimated Total Number of Errors is 0.1.
FEATURES
source 1. 160643
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="19"
/clone="CTC-492K19"
ORIGIN
Query Match 99.7%; Score 589.4; DB 8; Length 160643;
Best Local Similarity 99.8%; Pred. No. 1e-113;
Matches 590; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGAGGAGGAGGCTTGAAGAAGAACTCTGTTTGGAAAGAGAGAGAGGCTGGAG 60
DB 137453 ATGGTGGAGGCTTGAAGAAGAACTCTGTTTGGAAAGAGAGAGAGGCTGGAG 137394
QY 61 TGGAGTCCAGCAGGCTTCAAGACTACAGCAAGCCCTGCTCCGATCTCCCTAGCAAA 120
DB 137393 TGGAGTCCAGCAGGCTTCAAGACTACAGCAAGCCCTGCTCCGATCTCCCTAGCAAA 137334
QY 121 GTCCAGCGCAGCTTGGGCCCCCGAGACCCAGACCTCCGAGGAGATGCTCTCATCTAATAC 180
DB 137333 GTCCAGCGCAGCTTGGGCCCCCGAGACCCAGACCTCCGAGGAGATGCTCTCATCTAATAC 137274
QY 181 ACCCTTCAACAGCTTGAAGCTTGAAGCTTGGCTTCCGCCCCCTGCCCCCGAG 240
DB 137273 ACCCTTCAACAGCTTGAAGCTTGAAGCTTGGCTTCCGCCCCCTGCCCCCGAG 137214
QY 241 CCCCTCTTCTGGGAGAGAGATTTCTGCTGTGAGCCACATTTGGCTCTATCTCCAGG 300
DB 137213 CCCCTCTTCTGGGAGAGAGATTTCTGCTGTGAGCCACATTTGGCTCTATCTCCAGG 137154
QY 301 GAGCTGACACCTTCATGATGAGGAGTGAAGCCCTCAGAAATCCAGTGAATCTCCCTTGGC 360
DB 137153 GAGCTGACACCTTCATGATGAGGAGTGAAGCCCTCAGAAATCCAGTGAATCTCCCTTGGC 137094
QY 361 CTCGAGATGAAGTGCACCCAGGCTTATTCAGTCTTTTGAAGAGCTTGAAGCTTCCCGG 420
DB 137093 CTCGAGATGAAGTGCACCCAGGCTTATTCAGTCTTTTGAAGAGCTTGAAGCTTCCCGG 137034

QY 421 TACTTGGGAGACTTGGGCTTGATGATCTTTCTTGGACATTCGACATCTCGGTAGAA 480
DB 137033 TACTTGGGAGACTTGGGCTTGATGATCTTTCTTGGACATTCGACATCTCGGTAGAA 136974
QY 481 AAGGAGCTTGCAGCGGGCCCCACAGAGCTCTCTCAACCTCTTGGGCCAGGTTCT 540
DB 136973 AAGGAGCTTGCAGCGGGCCCCACAGAGCTCTCTCAACCTCTTGGGCCAGGTTCT 136914
QY 541 TGGAGTGAATGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 591
DB 136913 TGGAGTGAATGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 136863
RESULT 10
BD124690 825 bp DNA linear PART 18-SEP-2002
LOCUS
DEFINITION
ACCESSION BD124690.1 GI:23219635
VERSION BD124690.1 GI:23219635
KEYWORDS JP 2002017375-A/121.
SOURCE
ORGANISM Homo sapiens (human)
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
REFERENCE 1 (bases 1 to 825)
AUTHORS Ota,T., Nishikawa,T., Isegaki,T., Hayashi,K., Ishii,S., Kawai,Y.,
Wakamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and
Koga,H.
TITLE Primer for synthesizing full-length cDNA and use thereof
JOURNAL Patent: JP 2002017375-A 121 22-JAN-2002;
HELIX RESEARCH INSTITUTE
OS Homo sapiens (human)
PN JP 2002017375-A/121
PD 22-JAN-2002 JP 2002025172
PF 07-JUL-2000 JP 2002025172
PI TOSHIO OTA, TETSUO NISHIKAWA, TAKAO ISOGLI, KOJI HAYASHI, SHIZUKO
PI ISHII,
PI YURI KAWAI, AI WAKAMATSU, TOMOYASU SUGIYAMA, KEIICHI NAGAI, PI
SHINICHI KOJIMA,
PI TETSUJI OTSUKI, HISASHI KOGA
PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ 10,
PC C12P21/02, C12Q1/68, C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key
LOCATION/Qualifiers
FT source 1. 825
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 99.6%; Score 588.4; DB 6; Length 825;
Best Local Similarity 99.7%; Pred. No. 3.1e-113;
Matches 589; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGAGGAGGAGGCTTGAAGAAGAACTCTGATTTTGAAGAAGAGAGAGAGGTTGGAG 60
DB 93 ATGGTGGAGGCTTGAAGAAGAACTCTGATTTTGAAGAAGAGAGAGAGGTTGGAG 152
QY 61 TGGAGTCCAGCAGGCTTCAAGACTACAGCAAGCCCTGCTCCGATCTCCCTAGCAAA 120
DB 153 TGGAGTCCAGCAGGCTTCAAGACTACAGCAAGCCCTGCTCCGATCTCCCTAGCAAA 212
QY 121 GTCCAGCGCAGCTTGGGCCCCCGAGACCCAGCTCCGAGGAGATGCTCTCATCTAATAC 180
DB 213 GTCCAGCGCAGCTTGGGCCCCCGAGACCCAGCTCCGAGGAGATGCTCTCATCTAATAC 272
QY 181 ACCCTTCAACAGCTTGAAGCTTGAAGCTTGGCTTCCGCCCCCTGCCCCCGAG 240

Matches	589;	Conservative	0;	Mismatches	2;	Indels	0;	Gaps	0;
QY	1	ATGAGAGGAGGCTTGAAGAGAAAACCTCTGATTTTGGAAAGAGAGGAGGTGGAG	60						
Db	93	ATGATGGAGAGCTTGAAGAGAAAACCTCTGATTTTGAAGAGAGGAGGTGGAG	152						
QY	61	TGAGATCAGCAGGCTTTCAGAGCTACCAAGAGCCCTGCTCGCATCTCCCTAGACAA	120						
Db	153	TGAGATCAGCAGGCTTTCAGAGCTACCAAGAGCCCTGCTCGCATCTCCCTAGACAA	212						
QY	121	GTCCAGCGCAGCCTTGAGGCCCCCGAGCACCCAGCCTCCGAGGAGATGCTTCATCATAC	180						
Db	213	GTCCAGCGCAGCCTTGAGGCCCCCGAGCACCCAGCCTCCGAGGAGATGCTTCATCATAC	272						
QY	181	ACCTCTCAACAGCTGACAGCTGCACTTGGCTGGCTCCGCCCCCTGGCCCCCGAG	240						
Db	273	ACCTCTCAACAGCTGACAGCTGCACTTGGCTGGCTCCGCCCCCTGGCCCCCGAG	332						
QY	241	CCCCCTTCTTGAGGCGAGAGATTTCTCCCTGACAGCCACATTTGGCTATCTCAG	300						
Db	333	CCCCCTTCTTGAGGCGAGAGATTTCTCCCTGACAGCCACATTTGGCTATCTCAG	392						
QY	301	GAGCTGACACCTTCATGATGAGACTGAGCCCCCTCAGAAATCCAGTACTCCCCTTGGC	360						
Db	393	GAGCTGACACCTTCATGATGAGACTGAGCCCCCTCAGAAATCCAGTACTCCCCTTGGC	452						
QY	361	CTCCAGAAATGAATGCCACCCCAAGCCTGATTCAGTCTTTTGAAGAGCTGAGCTCCCG	420						
Db	453	CTCCAGAAATGAATGAAGGCCACCCCAAGCCTGATTCAGTCTTTTGAAGAGCTGAGCTCCCG	512						
QY	421	TACTTGAGGAGACTTGAGCTGAGATGATCTTCTTTCTGACATTTGACATGACATGCGGTGAA	480						
Db	513	TACTTGAGGAGACTTGAGCTGAGATGATCTTCTTTCTGACATTTGACATGACATGCGGTGAA	572						
QY	481	AAGGAGCCTGACAGGCGGCCACCAAGAGCCTCTTCACAACCTCTTGCCCCAGTTCT	540						
Db	573	AAGGAGCCTGACAGGCGGCCACCAAGAGCCTCTTCACAACCTCTTGCCCCAGTTCT	632						
QY	541	TGGAGTGGAAATGAATCGATCATCATGAGAAATCATTTCTGGGGTCTTAA	591						
Db	633	TGGAGTGGAAATGAATCGATCATCATGAGAAATCATTTCTGGGGTCTTAA	683						
RESULT 12									
LOCUS	CQ779981	825 bp	DNA	linear	PAT 17-MAR-2004				
DEFINITION	Sequence 121 from Patent EP1396543.								
ACCESSION	CQ779981								
VERSION	CQ779981.1	GI:45536053							
KEYWORDS									
SOURCE									
ORGANISM	Homo sapiens (human)								
REFERENCE	Homo sapiens								
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominiidae; Homo.								
TITLE									
JOURNAL	Oca,T., Nishikawa,T., Isogai,T., Hayashi,K., Ishii,S., Kawai,Y., Wakamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and Koga,H.								
FEATURES	Primers for synthesizing full length cDNA clones and their use Patent: EP 1396543-A 121 10-MAR-2004; Research Association for Biotechnology (JP)								
source	Location/Qualifiers 1..825 /organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606"								
ORIGIN									
Query Match	99.6%;	Score 588.4;	DB 6;	Length 825;					
Best Local Similarity	99.7%;	Pred. No. 3.1e-113;							
Matches	589;	Conservative	0;	Mismatches	2;	Indels	0;	Gaps	0;

QY 1 ATGAGGAGGCTTGAAGAGAAACACTCTGATTTTGAAGAGAGAGAGAGAGGTTGGAG 60
DB 93 ATGGGAGAGGCTTGAAGAGAAACACTCTGATTTTGAAGAGAGAGAGAGAGGTTGGAG 152
QY 61 TGGAGTCCAGAGAGGCTTCAAGAGCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
DB 153 TGGAGTCCAGAGAGGCTTCAAGAGCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 212
QY 121 GTCCAGCCAGAGCTTGGAG 180
DB 213 GTCCAGCCAGAGCTTGGAG 272
QY 181 ACCCTCCAAAGCTGAG 240
DB 273 ACCCTCCAAAGCTGAG 332
QY 241 CCCCTCTTCTGGAG 300
DB 333 CCCCTCTTCTGGAG 392
QY 301 GAGCTGACACCTCCATGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
DB 393 GAGCTGACACCTCCATGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 452
QY 361 CTCAGAAATGAAGTGCACACCCAGCTGATCCATTTTGAAGAGAGAGAGAGAGAGAG 420
DB 453 CTCAGAAATGAAGTGCACACCCAGCTGATCCATTTTGAAGAGAGAGAGAGAGAGAG 512
QY 421 TACTTGGAGAGAGCTGAG 480
DB 513 TACTTGGAGAGAGCTGAG 572
QY 481 AAGAGAGCTGACAG 540
DB 573 AAGAGAGCTGACAG 632
QY 541 TGGAGTGAATGAATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 591
DB 633 TGGAGTGAATGAATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 683

RESULT 13
LOCUS CQ781937 825 bp DNA linear PAT 17-MAR-2004
DEFINITION Sequence 2077 from Patent EP1396543.
ACCESSION CQ781937
VERSION CQ781937.1 GI:45537993
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE
AUTHORS Ota, T., Nishikawa, T., Isogai, T., Hayashi, K., Ishii, S., Kawai, Y.,
Wakamatsu, A., Sugiyama, T., Nagai, K., Kojima, S., Oseuki, T. and
Koga, H.
TITLE Primers for synthesizing full length cDNA clones and their use
JOURNAL Patent: EP 1396543-A 2077 10-MAR-2004;
Research Association for Biotechnology (JP)
FEATURES
source location/Qualifiers
1..825
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 99.6%; Score 588.4; DB 6; Length 825;
Best Local Similarity 99.7%; Pred. No. 3,1e-113;
Matches 589; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGAGGAGGCTTGAAGAGAAACACTCTGATTTTGAAGAGAGAGAGAGAGGTTGGAG 60
|||||

DB 93 ATGGTGGAGAGGCTTGAAGAGAAACACTCTGATTTTGAAGAGAGAGAGAGAGGTTGGAG 152
QY 61 TGGAGTCCAGAGAGGCTTCAAGAGCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
DB 153 TGGAGTCCAGAGAGGCTTCAAGAGCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 212
QY 121 GTCCAGCCAGAGCTTGGAG 180
DB 213 GTCCAGCCAGAGCTTGGAG 272
QY 181 ACCCTCCAAAGCTGAG 240
DB 273 ACCCTCCAAAGCTGAG 332
QY 241 CCCCTCTTCTGGAG 300
DB 333 CCCCTCTTCTGGAG 392
QY 301 GAGCTGACACCTCCATGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
DB 393 GAGCTGACACCTCCATGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 452
QY 361 CTCAGAAATGAAGTGCACACCCAGCTGATCCATTTTGAAGAGAGAGAGAGAGAGAG 420
DB 453 CTCAGAAATGAAGTGCACACCCAGCTGATCCATTTTGAAGAGAGAGAGAGAGAGAG 512
QY 421 TACTTGGAGAGAGCTGAG 480
DB 513 TACTTGGAGAGAGCTGAG 572
QY 481 AAGAGAGCTGACAG 540
DB 573 AAGAGAGCTGACAG 632
QY 541 TGGAGTGAATGAATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 591
DB 633 TGGAGTGAATGAATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 683

RESULT 14
LOCUS AC021625 164652 bp DNA linear HTG 28-MAR-2000
DEFINITION Homo sapiens clone RP11-384B6, WORKING DRAFT SEQUENCE, 34 unordered
pieces.
ACCESSION AC021625
VERSION AC021625.2 GI:7331453
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE
AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.
TITLE 1 (bases 1 to 164652)
JOURNAL Birren, B., Linton, L., Nusbaum, C. and Lander, E.
REFERENCE
AUTHORS 2 (bases 1 to 164652)
unpublished
Birren, B., Linton, L., Nusbaum, C. and Lander, E.
Anderson, S., Baldwin, J., Barna, N., Beckert, R., Beda, F.,
Boguslavsky, L., Bouhagalter, B., Brown, A., Burkett, G., Castle, A.,
Chepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P.,
DeRellano, K., Dewar, K., Domino, M., Doyle, M., Fensholt, J.,
Ferreira, P., Fitzhugh, W., Forrest, C., Gage, D., Galagan, J.,
Gardyna, S., Grant, G., Hago, B., Heald, A., Horton, L.,
Howland, J. C., Johnson, R., Jones, C., Kahn, L., Karatas, A., Klein, J.,
Landers, T., Lehotzky, J., Levine, R., Liu, C., Liu, G., Locke, K.,
Macdonald, P., Marquis, N., McEwan, P., McGuirk, A., McKernan, K.,
McNeel, R., Menzies, J., Menzies, L., Morrow, J., Naylor, J.,
Norman, C. H., O'Connor, T., O'Donnell, P., O'Leary, T. M., Peterson, K.,
Pierre, N., Pisan, C., Pollara, V., Raymond, C., Riley, R., Rothman, D.,
Roy, A., Santos, R., Severy, P., Spencer, B., Stange-Thomann, N.,
Stojanovic, N., Subramanian, A., Talamas, J., Teesdale, S., Theodore, J.,
Tirrell, A., Vassiliev, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W. J.,

TITLE
JOURNAL
COMMENT

Zimmer, A. and Zody, M.

Direct Submission

Submitted (16-JAN-2000) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

On Mar 28, 2000 this sequence version replaced gi:6705474.

All repeats were identified using RepeatMasker:

Smith, A.P. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: MIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

Project Information

Center project name: L5880

Center clone name: 384_B_6

Summary Statistics

Sequencing vector: M13; M7815; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 145359 bases at least Q40

Consensus quality: 155420 bases at least Q30

Consensus quality: 159351 bases at least Q20

Insert size: 161352; sum-of-contigs

Quality coverage: 4.0 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 34 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 1083: contig of 1083 bp in length
1084 1183: gap of 100 bp
1184 2479: contig of 1286 bp in length
2480 2579: gap of 100 bp
2580 4311: contig of 1732 bp in length
4312 4411: gap of 100 bp
4411 5951: contig of 1540 bp in length
5952 6051: gap of 100 bp
6052 7761: contig of 1710 bp in length
7762 7861: gap of 100 bp
7862 9630: contig of 1769 bp in length
9631 9731: gap of 100 bp
9731 11763: contig of 2033 bp in length
11764 11863: gap of 100 bp
11864 14294: contig of 2431 bp in length
14295 14394: gap of 100 bp
14395 16731: contig of 2337 bp in length
16732 16831: gap of 100 bp
16832 20279: contig of 3448 bp in length
20280 20379: gap of 100 bp
20380 23751: contig of 3372 bp in length
23752 23851: gap of 100 bp
23852 26903: contig of 3052 bp in length
26904 27003: gap of 100 bp
27004 31228: contig of 4225 bp in length
31229 31328: gap of 100 bp
31329 35336: contig of 4008 bp in length
35337 35436: gap of 100 bp
35437 38469: contig of 3033 bp in length
38470 38569: gap of 100 bp
38570 42543: contig of 4074 bp in length
42544 42743: gap of 100 bp
42744 47087: contig of 4334 bp in length
47088 47187: gap of 100 bp
47188 51475: contig of 4288 bp in length
51476 51575: gap of 100 bp
51576 56093: contig of 4558 bp in length
56094 56193: gap of 100 bp
56194 59387: contig of 3194 bp in length

FEATURES

SOURCE

* 59388 59487: gap of 100 bp
* 59488 63987: contig of 4300 bp in length
* 63988 64087: gap of 100 bp
* 64088 69347: contig of 5260 bp in length
* 69348 69447: gap of 100 bp
* 69448 74268: contig of 4821 bp in length
* 74269 74369: gap of 100 bp
* 74369 79365: contig of 4997 bp in length
* 79366 79465: gap of 100 bp
* 79466 83211: contig of 3746 bp in length
* 83212 83311: gap of 100 bp
* 83312 88101: contig of 4790 bp in length
* 88102 88201: gap of 100 bp
* 88202 93785: contig of 5584 bp in length
* 93786 93885: gap of 100 bp
* 93886 100407: contig of 6522 bp in length
* 100408 100507: gap of 100 bp
* 100508 108158: contig of 7651 bp in length
* 108159 108259: gap of 100 bp
* 108259 115631: contig of 7373 bp in length
* 115632 115731: gap of 100 bp
* 115732 125488: contig of 9757 bp in length
* 125489 125589: gap of 100 bp
* 125589 136013: contig of 10425 bp in length
* 136014 136113: gap of 100 bp
* 136114 149110: contig of 12997 bp in length
* 149111 149210: gap of 100 bp
* 149211 164652: contig of 15442 bp in length.

Location/Qualifiers

1..164652

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/clone="RP11-384B6"

/clone_1lb="RPC1-11 Human Male BAC"

1..1083

/note="assembly_fragment"

1084..1183

/estimated_length=100

1184..2479

/note="assembly_fragment"

2480..2579

/estimated_length=100

2580..4311

/note="assembly_fragment"

4312..4411

/estimated_length=100

4412..5951

/note="assembly_fragment"

5952..6051

/estimated_length=100

6052..7761

/note="assembly_fragment"

7762..7861

/estimated_length=100

7862..9630

/note="assembly_fragment"

9631..9730

/estimated_length=100

9731..11763

/note="assembly_fragment"

11764..11863

/estimated_length=100

11864..14294

/note="assembly_fragment"

14295..14394

/estimated_length=100

14395..16731

/note="assembly_fragment"

16732..16831

/estimated_length=100

16832..20279

/note="assembly_fragment"

```
gap          20280..20379
             /estimated_length=100
misc_feature 20380..23751
             /note="assembly_fragment"
gap          23752..23851
             /estimated_length=100
misc_feature 23852..26903
             /note="assembly_fragment"
gap          26904..27003
             /estimated_length=100
misc_feature 27004..31228
             /note="assembly_fragment"
gap          31229..31328
             /estimated_length=100
misc_feature 31329..35336
             /note="assembly_fragment"
gap          35337..35436
             /estimated_length=100
misc_feature 35437..38469
             /note="assembly_fragment"
gap          38470..38569
             /estimated_length=100
```

Query Match 98.9%; Score 584.6; DB 14; Length 164652;
Best Local Similarity 99.3%; Pred. No. 1e-112;
Matches 587; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
QY 1 ATGAGGAGGAGGCTTGAAGAGAAACCTGTGATTTGGAAGAGAGAGAGAGGCTGGAG 60
DB 113509 ATGGTGGAGGCTTGAAGAGAAACCTGTGATTTGGAAGAGAGAGAGAGGCTGGAG 113568
QY 61 TGGAGTCCAGAGGCTTGAAGAGAAACCTGTGATTTGGAAGAGAGAGAGAGGCTGGAG 120
DB 113569 TGGAGTCCAGAGGCTTGAAGAGAAACCTGTGATTTGGAAGAGAGAGAGAGGCTGGAG 113628
QY 121 GTCCAGCGAGGCTTGAAGAGAAACCTGTGATTTGGAAGAGAGAGAGAGGCTGGAG 180
DB 113629 GTCCAGCGAGGCTTGAAGAGAAACCTGTGATTTGGAAGAGAGAGAGAGGCTGGAG 113688
QY 181 ACCCTCCAGAGGCTTGAAGAGAAACCTGTGATTTGGAAGAGAGAGAGAGGCTGGAG 240
DB 113689 ACCCTCCAGAGGCTTGAAGAGAAACCTGTGATTTGGAAGAGAGAGAGAGGCTGGAG 113748
QY 241 CCCCTCTTCTGGGAGAGAGATTTCTCCCTGTGAGAGAGAGAGAGAGGCTGGAG 300
DB 113749 CCCCTCTTCTGGGAGAGAGATTTCTCCCTGTGAGAGAGAGAGAGAGGCTGGAG 113808
QY 301 GAGCTGACACCTCATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGGCTGGAG 360
DB 113809 GAGCTGACACCTCATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGGCTGGAG 113868
QY 361 CTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGGCTGGAG 420
DB 113869 CTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGGCTGGAG 113928
QY 421 TACTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGGCTGGAG 480
DB 113929 TACTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGGCTGGAG 113988
QY 481 AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGGCTGGAG 540
DB 113989 AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGGCTGGAG 114048
QY 541 TGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGGCTGGAG 591
DB 114049 TGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGGCTGGAG 114099
```

RESULT 15
CQ463286 CQ463286 506 bp DNA linear PAT 30-JAN-2004
LOCUS
DEFINITION Sequence 6064 from Patent WO0192581.
ACCESSION CQ463286

VERSION CQ463286.1 GI:41428905
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE
1 Algate, P.A., Harlocker, S.L., and Jones, R.
Compositions and methods for the therapy and diagnosis of
ovarian cancer
Patent: WO 0192581-A 6064 06-DEC-2001;
JOURNAL CORIXA CORPORATION (US)
FEATURES
source
1..506
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

Query Match 80.0%; Score 473; DB 6; Length 506;
Best Local Similarity 98.6%; Pred. No. 5.3e-89;
Matches 488; Conservative 0; Mismatches 5; Indels 2; Gaps 1;

```
QY 60 GTGAGTCCAGAGGCTTGAAGAGAAACCTGTGATTTGGAAGAGAGAGAGAGGCTGGAG 119
DB 14 GCGAGTCCAGAGGCTTGAAGAGAAACCTGTGATTTGGAAGAGAGAGAGAGGCTGGAG 73
QY 120 AGTCCAGAGGAGAGGCTTGAAGAGAAACCTGTGATTTGGAAGAGAGAGAGAGGCTGGAG 179
DB 74 AGTCCAGAGGAGAGGCTTGAAGAGAAACCTGTGATTTGGAAGAGAGAGAGAGGCTGGAG 131
QY 180 CACCTCCAGAGGAGAGGCTTGAAGAGAAACCTGTGATTTGGAAGAGAGAGAGAGGCTGGAG 239
DB 132 CACCTCCAGAGGAGAGGCTTGAAGAGAAACCTGTGATTTGGAAGAGAGAGAGAGGCTGGAG 191
QY 240 GCCCTCTTCTGGGAGAGAGATTTCTCCCTGTGAGAGAGAGAGAGAGGCTGGAG 299
DB 192 GCCCTCTTCTGGGAGAGAGATTTCTCCCTGTGAGAGAGAGAGAGAGGCTGGAG 251
QY 300 GAGCTGACACCTCATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGGCTGGAG 359
DB 252 GAGCTGACACCTCATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGGCTGGAG 311
QY 360 CTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGGCTGGAG 419
DB 312 CTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGGCTGGAG 371
QY 420 GTAATTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGGCTGGAG 479
DB 372 GTAATTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGGCTGGAG 431
QY 480 AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGGCTGGAG 539
DB 432 AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGGCTGGAG 491
QY 540 TTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGGCTGGAG 591
DB 492 TTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGGCTGGAG 540
```

Search completed: February 6, 2006, 11:34:39
Job time : 3211 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 6, 2006, 10:26:36 ; Search time 476 Seconds
(without alignments)
8274.860 Million cell updates/sec

Title: US-10-069-386A-1

Perfect score: 591
Sequence: 1 atggagggagagctgaagag.....aatcatctcggggtcctaa 591

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_21.*
1: geneeqn1980s:*
2: geneeqn1990s:*
3: geneeqn2000s:*
4: geneeqn2001as:*
5: geneeqn2001bs:*
6: geneeqn2002as:*
7: geneeqn2002bs:*
8: geneeqn2003as:*
9: geneeqn2003bs:*
10: geneeqn2003cs:*
11: geneeqn2003ds:*
12: geneeqn2004as:*
13: geneeqn2004bs:*
14: geneeqn2005s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	591	100.0	591	4	AAf28052 Replicat
2	589.4	99.7	1325	4	AAf28052 Human ful
3	589.4	99.7	1325	12	ADL32051 Full leng
4	589.4	99.7	1515	13	ADQ85787 Human tum
5	589.4	99.7	1542	10	ADG33127 Human DNA
6	589.4	99.7	1542	14	ADZ70575 Human CDN
7	588.4	99.6	825	4	AAK91661 Human CDN
8	588.4	99.6	825	4	AAK93617 Human CDN
9	588.4	99.6	825	12	ADL30044 3' end of
10	588.4	99.6	825	12	ADL28088 5' end of
11	578.4	97.9	1596	4	AAH34147 Human gen
12	576.2	97.5	1596	4	AAH34147 Human col
13	473	80.0	506	6	ABJ83086 Human ova
14	186.6	31.6	205	6	ABJ83261 Human ova
15	166	28.1	526	12	ACH77983 Human gen
16	47.4	8.0	2082	6	AAJ41497 Mouse nae
17	46.4	7.9	51552	6	AAJ96607 DNA encod
18	42.6	7.2	1603	8	ACC51095 Human Sho
19	42.4	7.2	511	12	ACH91786 Human gen

20	42.4	7.2	600	6	ABQ52497	ABq52497 Oligonuc
21	42.4	7.2	600	6	ABQ52496	ABq52496 Oligonuc
22	42.4	7.2	1850	8	ADA70325	Ada70325 Rice gene
23	41.8	7.1	32329	12	ADQ51695	Adq51695 Streptomy
24	41.4	7.0	825	12	ADQ00165	Adq00165 Novel hum
25	41.4	7.0	825	12	ADN98596	Adn98596 Novel hum
26	41	6.9	1925	2	AAK90924	Aak90924 Epretein B
27	41	6.9	1926	3	AAK50254	Aak50254 Epretein B
28	41	6.9	1926	4	AAK82902	Aak82902 EBV tethe
29	41	6.9	1926	10	ADK65580	Adk65580 Human her
30	41	6.9	1926	14	ADK68155	Adk68155 Epretein-B
31	41	6.9	2046	14	ABE68224	Abe68224 Epretein-B
32	41	6.9	2580	3	AAK75454	Aak75454 Nucleotid
33	41	6.9	2580	6	AAI64275	AAi64275 Epretein-B
34	41	6.9	5452	2	AAK90923	AAk90923 Anti-sens
35	41	6.9	8705	2	AAZ23778	Aaz23778 Vector PS
36	41	6.9	8705	12	ADM10659	Adm10659 Expressio
37	41	6.9	9482	12	ADP64415	Adp64415 Vector PC
38	41	6.9	9482	14	ADW05161	Adw05161 S/MAR vec
39	41	6.9	9600	2	AAV21683	Aav21683 Vector p1
40	41	6.9	10285	6	ABG71027	Abg71027 PCBP-Xa-P
41	41	6.9	10285	6	ABG66453	Abg66453 Plasmid p
42	41	6.9	10330	12	ADL67154	Adl67154 Plasmid p
43	41	6.9	10380	2	AAZ22248	Aaz22248 Nucleotid
44	41	6.9	10477	12	ADL67152	Adl67152 Plasmid p
45	41	6.9	10516	12	ADL67150	Adl67150 Plasmid p

ALIGNMENTS

RESULT 1	
AAf28052	AAf28052 standard; DNA; 591 BP.
ID	AAf28052; (first entry)
AC	23-MAY-2001
XX	Replication protein A binding transcriptional activator 1 RBT1 gene.
XX	RBT1; replication protein A binding transcriptional activator 1; RPA32;
KW	gene therapy; apoptosis; cancer; leukaemia; ds.
XX	Unidentified.
OS	
XX	Key
FH	Location/Qualifiers
FT	1..591
FT	/*tag= a
FT	/product= "RBT1"
XX	WO200114546-A2.
XX	01-MAR-2001.
PD	17-AUG-2000; 2000WO-CA000948.
XX	19-AUG-1999; 99US-0149472P.
PR	(TRAN-) CENT TRANSLATIONAL RES IN CANCER.
PA	Alaoui-Jamali MA, Cho JM;
XX	WPI; 2001-218447/22.
PI	P-PSDB; AAB35402.
XX	Novel replication protein A binding transcriptional activator 1 gene,
PT	useful for treating neoplastic disorders such as cancer and in gene
PT	therapy.
XX	Claim 1; Fig 1; 16pp; English.
PS	The present invention provides the protein and coding sequences of the
XX	
CC	

Dd		633	IGGAGTGGATGAACGTGATCATCATGTAGAAATCATTCTGGGTCCTTA	683
XX	RESULT 3			
XX	ADL32051			
XX	ID	ADL32051	standard; cDNA; 1325 BP.	
XX	AC	ADL32051;		
XX	DT	20-MAY-2004	(first entry)	
XX	DE	Full length human cDNA clone Segid 4084.		
XX	KW	human; medicine; signal transduction; glycoprotein; transcription; oligo-capping method; ss; gene.		
XX	OS	Homo sapiens.		
XX	FN	EP1396543-A2.		
XX	PD	10-MAR-2004.		
XX	PE	07-JUL-2000; 2003BP-00025638.		
XX	PR	08-JUL-1999; 99JP-00134486.		
XX	PR	11-JAN-2000; 2000JP-00118774.		
XX	PR	02-MAY-2000; 2000JP-00183865.		
XX	PR	07-JUL-2000; 2000EP-00114089.		
XX	PA	(REAS-) RES ASSOC BIOTECHNOLOGY.		
XX	P1	Ota T, Nishikawa T, Isogai T, Hayashi K, Iehli S, Kawai Y,		
XX	P1	Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;		
XX	DR	WPI; 2004-204755/20.		
XX	DR	P-PsDB; ADL32052.		
XX	PT	New oligonucleotide primers (830 CDNA) useful for synthesizing full length human cDNAs.		
XX	PS	Example 1; SEQ ID NO 4084; 1340bp; English.		
XX	CC	This invention relates to a novel primers useful for synthesizing full length cDNA molecules that encode human proteins. Specifically, it refers to secretory or membrane proteins that are potential therapeutic agents/ target molecules in the field of medicine, and in particular genes encoding proteins that are associated with signal transduction, glycoproteins and transcription. The present invention describes a method for efficiently cloning a full length human cDNA from both the 5' and 3' ends using the oligo-capping method. This polynucleotide sequence is a full length human cDNA clone of the invention.		
XX	SO	Sequence 1325 BP; 263 A; 371 C; 359 G; 332 T; 0 U; 0 Other;		
Qy	Query Match	99.7%; Score 589.4; DB 12; Length 1325;		
Dd	Best Local Similarity	99.8%; Pred. No. 1,4e-134;		
XX	Matches	590; Conservative 0; Mismatches 1; Indels 0; Gaps 0		
Qy	1	ATGAGGAGGAGCGTTTGAAAGAGAAAACACTTGTATTTGGAAAGAGAGAGAGGCGGGAG	60	
Dd	93	ATGGTGGAGAGGGTTTGAAAGAGAAAACACTCTGATTTGGAAAGAGAGAGAGGCGGGAG	152	
Qy	61	TGAGATCCAGCAGGCGCTTCAGAGCTAACGAGAAGGCCCTGCCTCCCTTAGACANA	120	
Dd	153	TGAGATCCAGCAGGCGCTTCAGAGCTAACGAGAAGGCCCTGCCTCCCTTAGACANA	212	
Qy	121	GTCACGCGAAGCTGGGCCCCCGAGACACCAGCCTTCGAGGCAATGCTCATTCATAAC	180	
Dd	213	GTCACGCGAAGCTGGGCCCCCGAGACACCAGCCTTCGAGGCAATGCTCATTCATAAC	272	
Qy	181	ACCTTCGAACAGCTGCAAGCTGCACTTGGCTCTCCGCCCCCTGCTGCCCCCGAG	240	
Dd	273	ACCTTCGAACAGCTGCAAGCTGCACTTGGCTCTCCGCCCCCTGCTGCCCCCGAG	332	

QY	241	CCCCCTCTTCTGGGCGAGAGAGATTTCCTCCTGTACGCCACCATTTGGCTTATCTCTCAGG	300
Db	333	CCCCCTCTTCTGGGCGAGAGAGATTTCCTCCTGTACGCCACCATTTGGCTTATCTCTCAGG	392
QY	301	GAGCTGGACACCTCCATGGATTGGGACTGAGCCCCCAGAAATCAGTGACTCCCTTGGC	360
Db	393	GAGCTGGACACCTCCATGGATTGGGACTGAGCCCCCAGAAATCAGTGACTCCCTTGGC	452
QY	361	CTCCAGAAATGAATGACCAACCCAGCCTGATCCAGTCTTTTGAAGCCTTGAGCTCCCG	420
Db	453	CTCCAGAAATGAATGACCAACCCAGCCTGATCCAGTCTTTTGAAGCCTTGAGCTCCCG	512
QY	421	TACTTTGGGGGAGACTCTGCGCTGTGATGACTTTCTTTGGAATTTGACACATCTGGGGTAGAA	480
Db	513	TACTTTGGGGGAGACTCTGCGCTGTGATGACTTTCTTTGGAATTTGACACATCTGGGGTAGAA	572
QY	481	AAGGAGCCTGACCGGGGCCCCACAGAGCCTCTCTCAACACTCTTGTGTCGCCAGGTTCT	540
Db	573	AAGGAGCCTGACCGGGGCCCCACAGAGCCTCTCTCAACACTCTTGTGTCGCCAGGTTCT	632
QY	541	TGGGAGTGAATGAATGATCGATCATCATGGAATCATTTCTGGGCTCTTAA	591
Db	633	TGGGAGTGAATGAATGATCGATCATCATGGAATCATTTCTGGGCTCTTAA	683
RESULT 4			
ID	ADQ85787	standard; cDNA; 1515 BP.	
XX	ADQ85787;		
XX	ADQ85787;		
DT	07-OCT-2004	(first entry)	
DE	Human tumour-associated antigenic target (TAT) cDNA sequence #2601.		
XX	human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;		
KW	cancer; cell proliferative disorder; gene; ss.		
XX	Homo sapiens.		
OS	Homo sapiens.		
PN	WO2004060270-A2.		
PD	22-JUL-2004.		
XX	15-OCT-2003; 2003WO-US029126.		
PF	15-OCT-2003; 2003WO-US029126.		
XX	18-OCT-2002; 2002US-0418988P.		
PR	18-OCT-2002; 2002US-0418988P.		
XX	(GETH) GENENTECH INC.		
PA	(WUTD/) WU T D.		
PA	(ZHOU/) ZHOU Y.		
PI	Wu TD, Zhou Y;		
XX	WPI, 2004-534300/51.		
DR	WPI, 2004-534300/51.		
XX	New nucleic acid molecule and encoded polypeptide, for diagnosing,		
PT	preventing or treating cell proliferative disorders such as cancer.		
PS	Claim 1; SEQ ID NO 2601; 5504bp; English.		
XX	Claim 1; SEQ ID NO 2601; 5504bp; English.		
CC	The present invention describes an isolated tumour-associated antigenic		
CC	target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide		
CC	sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of		
CC	(a); (c) the complement of (a) or (b); (d) a sequence that has 80%		
CC	sequence identity to (a)-(c); or (e) a sequence that hybridizes to (a)-		
CC	(c). Also described: (1) an expression vector comprising the above		
CC	nucleic acid; (2) a host cell comprising the above expression vector; (3)		
CC	a process for producing a polypeptide; (4) an isolated polypeptide		
CC	comprising: (a) an amino acid sequence encoded by any of the above		
CC	nucleotide sequences; (b) an amino acid sequence encoded by the full-		
CC	length coding region of the above nucleotide sequences; or (c) a sequence		

CC having at least 80% identical to (a) or (b); (5) a chimeric polypeptide
CC comprising the above polypeptide fused to a heterologous polypeptide; (6)
CC an isolated antibody that binds to the above polypeptide; (7) a process
CC for producing the antibody; (8) an isolated oligopeptide that binds to
CC the above polypeptide; (9) a tumour-associated antigenic target (TAT)
CC binding organic molecule that binds to the above polypeptide; (10) a
CC composition of matter comprising the above (chimeric) polypeptide,
CC antibody, oligopeptide or TAT binding organic molecule, in combination
CC with a carrier; (11) an article of manufacture comprising a container and
CC the composition of matter contained within the container; (12) methods of
CC inhibiting the growth of a cell that expresses the above protein, where
CC the growth of the cell is at least in part dependent upon a growth
CC potentiating effect of the above protein; (13) a method of
CC therapeutically treating a mammal having a cancerous tumour comprising
CC cells that express the above protein; (14) a method of determining the
CC presence of a protein in a sample suspected of containing the protein
CC described above; (15) methods of diagnosing the presence of a tumour in a
CC mammal; (16) a method for treating or preventing a cell proliferative
CC disorder associated with increased expression or activity of the above
CC protein; and (17) a method of binding an antibody, oligopeptide or
CC organic molecule to a cell that expresses the protein described above.
CC The TAT sequences have cytostatic activities, and can be used in gene
CC therapy. The composition and methods are useful for diagnosing,
CC preventing or treating cancer. The composition is also used for preparing
CC a medicament for the therapeutic treatment or diagnostic detection of a
CC cell proliferative disorder or cancer. The present sequence represents a
CC human TAT cDNA sequence from the present invention.

Sequence 1515 BP; 328 A; 404 C; 385 G; 398 T; 0 U; 0 Other;

Query Match	99.7%;	Score 589.4;	DB 13;	Length 1515;
Best Local Similarity	99.8%;	Pred. No. 1.5e-134;		

Matches 590; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	1	ATGAGAGGAGGCTTGAAGAAGAAACACTCTGATTTTGGAAAGAGAGAGAGAGAGTGGAG	60
Db	279	ATGGTGGAGGGCTTGAAGAAGAAACACTGATTTTGAAGAAGAGAGAGAGTGGAG	358
QY	61	TGAGGTCGACAGGCTTTCAGAGCTTCCAGCAAGCCTGCTCCGCACTTCCTCAGACAA	120
Db	339	TGAGATTCAGCAGGCTTTCAGAGCTTCCAGCAAGCCTGCTCCGCACTTCCTCAGACAA	398
QY	121	GTCCAGCGCAGCCTTGGGCCCCCGAGCACCAGCTCCGACAGGAGATGTCTGATCATTAAC	180
Db	399	GTCCAGCGCAGCCTTGGGCCCCCGAGCACCAGCCTCCGACAGGAGATGTCTGATCATTAAC	458
QY	181	ACCTTCACACAGCTGAGAGCTGCACTTGGCTTCGGCCCTTCGCTCCGCCCCCGAG	240
Db	459	ACCTTCACACAGCTGAGAGCTGCACTTGGCTTCGGCCCTTCGCTCCGCCCCCGAG	518
QY	241	CCCTCTCTCTGGGCGAGAGGATTTCTCCCTGTACGCCACATTTGGCTTATCTCAGG	300
Db	519	CCCTCTCTCTGGGCGAGAGGATTTCTCCCTGTACGCCACATTTGGCTTATCTCAGG	578
QY	301	GAGCTGACACCTCCATGAGATGAGAGCTGAGCCCCCTCAGAAATCAGTGACTCCCTTGGC	360
Db	579	GAGCTGACACCTCCATGAGATGAGAGCTGAGCCCCCTCAGAAATCAGTGACTCCCTTGGC	638
QY	361	CTCCAGATGAAGATGTCACCCAGAGCTGATCCAGTCTTCTTAAAGCTCTGAACTCCCG	420
Db	639	CTCCAGATGAAGATGTCACCCAGAGCTGATCCAGTCTTCTTAAAGCTCTGAACTCCCG	698
QY	421	TACTTGGGGGAGCTCTGAGCTGATGATCTTCTTGTGGACATTTGACATCTGCGGTGAA	480
Db	699	TACTTGGGGGAGCTCTGAGCTGATGATCTTCTTGTGGACATTTGACATCTGCGGTGAA	758
QY	481	AAGAGGCTTGACAGGAGCCGCCACAGAGCCTCTCAACAACCTCTTGTGTGCCCAAGTTCT	540
Db	759	AAGAGGCTTGACAGGAGCCGCCACAGAGCCTCTCAACAACCTCTTGTGTGCCCAAGTTCT	818
QY	541	TGGAGATGGAATGAATCTGATTCACATATGGAATCATTTCTGGGCTCTAA	591
Db	819	TGGAGATGGAATGAATCTGATTCACATATGGAATCATTTCTGGGCTCTAA	869

RESULT 5
ADG33127
ID ADG33127 standard; DNA; 1542 BP.

26-FBB-2004 (first entry)
Human DNA differentially expressed in patients with SLR SeqID451

human; ds; autoimmune; chronic inflammatory disease; SLE; systemic lupus erythematosus; rheumatoid arthritis; cholecystitis; Sjogren's disease; CRST syndrome; sclerodema; ankylosing spondylitis; ulcerative colitis; primary sclerosing cholangitis; appendicitis; diverticulitis; primary biliary sclerosis.

Homo sapiens.

WO2003090694-A2

06-NOV-2003

24-APR-2003; 2003WO-US013015.

24-APR-2002; 2002US-00131827.

(EXPR-) EXPRESSION DIAGNOSTICS INC.

Wohlgemuth J, Fry K, Woodward R, Ly N;

WPI; 2003-877243/81.

Diagnosing or monitoring autoimmune and chronic inflammatory diseases, such as rheumatoid arthritis, systemic lupus erythematosus, ulcerative colitis, psoriasis and asthma by detecting the expression level of one or more genes.

Claim 18; SEQ ID NO 451; 877pp; English.

This invention relates to novel methods for diagnosing and monitoring autoimmune and chronic inflammatory diseases. Specifically, it refers to the identification of genes that have a clinical utility as diagnostic tools for the management of, in particular, patients with systemic lupus erythematosus (SLE) or rheumatoid arthritis (RA). Accordingly, the present invention describes a method for determining the levels of multiple differentially expressed genes of a patient, in a concerted manner, in order to achieve an improved diagnostic assay with sensitivity and specificity for the disease in question. As such, these genes are useful for the diagnosis of various other inflammatory disorders including cholecystitis, Sjogren's disease, CRST syndrome, scleroderma, ankylosing spondylitis, ulcerative colitis, primary sclerosing cholangitis, appendicitis, diverticulitis, and primary biliary sclerosis. This polynucleotide is a DNA sequence representing human mRNA that is differentially expressed in patients with SLE, used in an exemplification of the invention.

SQ Sequence 1542 BP; 346 A; 405 C; 388 G; 403 T; 0 U; 0 Other;

Query Match	99.7%;	Score 589.4;	DB 10;	Length 1542;
Best Local Similarity	99.8%;	Pred. No. 1.5e-134;		

Matches 590; Conservative 0; Mismatches 1; Indels 0; Gaps 0

QY 1 ATGAGAGGAGGCTTTGAAAGGAAACCTCTGATTTTGAAAGAGAGAGAGAGTGGAG 60
Db 292 ATGGTGGAGGGGCTTGAAGAGGAAACCTCTGATTTTGAAAGAGAGAGAGAGTGGAG 355
QY 61 TCGAGTCCAGCAGGCGCTTCAAGACTACCAAGCAAGCGCTCTCGGATTCCTCCAGACAA 120
Db 352 TCGAGTCCAGCAGGCGCTTCAAGACTACCAAGCAAGCGCTCTCGGATTCCTCCAGACAA 411
QY 121 GTTCAGCGCAGGCTGGGGCCCCCGAGCAGCAGCCAGCTCTCGCAGGAGATGTCTCATTCATTAAC 180

```
Db 412 GTCCAGCGAGCTGGGCCCCGAGAGACCCAGCCTCCGAGGACATGTCTTCATTCATAC 471
Qy 181 ACCCTCAACAGCTGAGGCTGACTTGGCTGGCTCCGCTCCGCTGCCCCCGAG 240
Db 472 ACCCTCAACAGCTGAGGCTGACTTGGCTGGCTCCGCTCCGCTGCCCCCGAG 531
Qy 241 CCCCTCTTCTGGGCGAGGAGATTTCTCCCTGTGAGCCACCATTTGCTCTACAG 300
Db 532 CCCCTCTTCTGGGCGAGGAGATTTCTCCCTGTGAGCCACCATTTGCTCTACAG 591
Qy 301 GAGCTGACACCTCCATGATGAGTGAAGCCCTCAGAAATCCAGTACTCCCTTGGC 360
Db 592 GAGCTGACACCTCCATGATGAGTGAAGCCCTCAGAAATCCAGTACTCCCTTGGC 651
Qy 361 CTCGAGATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 420
Db 652 CTCGAGATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 711
Qy 421 TACTTGGGGGACTCTGGCTGTGATGATCTTCTTGTGACATTGACATCTGCGTAGA 480
Db 712 TACTTGGGGGACTCTGGCTGTGATGATCTTCTTGTGACATTGACATCTGCGTAGA 771
Qy 481 AAGAGCCTGACCGGGCCCCCAGCAGGCTCTCAGAACCTTCTTGTGCGCCAGTTCT 540
Db 772 AAGAGCCTGACCGGGCCCCCAGCAGGCTCTCTCAACACTTCTGTGCGCCAGTTCT 831
Qy 541 TGGGAGTGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 591
Db 832 TGGGAGTGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 882
```

RESULT 6

ADZ70575 standard; cDNA; 1542 BP.

ADZ70575;

30-JUN-2005 (first entry)

Human cDNA from lung cancer marker gene RBP1.

Tumor marker; ss; gene; lung tumor; cytosolic; neoplasm; expression; DNA microarray.

Homo sapiens.

WO2005032495-A2.

14-APR-2005.

01-OCT-2004; 2004WO-US034163.

03-OCT-2003; 2003US-0508355P.

(PARB) BAYER PHARM CORP.

Taylor I, Pauloeki NR, Bigwood D;

WPI; 2005-285325/29.

P-PSDB; ADZ70576.

Providing a patient diagnosis for lung cancer comprises comparing the level of expression of genes or gene products in a biological sample from the patient with that from a normal individual.

Claim 2; SEQ ID NO 260; 60pp; English.

The invention relates to providing a patient diagnosis for lung cancer comprising comparing the level of expression of genes or gene products in a biological sample from the patient with the level of expression of genes or gene products in a biological sample from a normal individual. Also included are distinguishing between normal and disease tissues,

monitoring the response of a patient being treated for lung cancer by administering an anti-cancer agent, identifying a compound useful for the treatment of lung cancer and an array for distinguishing between normal and disease tissues (comprising 2 or more probes corresponding to 2 or more genes selected from any of the 200 nucleotide sequences given in the specification, or 2 or more polypeptides comprising any of the 200 amino acid sequences given in the specification). In providing a patient diagnosis for lung cancer, one or more genes are selected from any of the 200 nucleotide sequences as mentioned in the specification, or one or more gene products are polypeptides selected from any of the 20 amino acid sequences mentioned in the specification. The methods are useful for detecting and treating lung cancer. These may also be used for designing, identifying and optimizing therapeutics for cancer. The present sequence represents a cDNA from one of the 200 lung cancer marker genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 1542 BP; 346 A; 405 C; 388 G; 403 T; 0 U; 0 Other;

Query Match 99.7%; Score 589.4; DB 14; Length 1542;

Best Local Similarity 99.8%; Pred. No. 1.5e-134;

Matches 590; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy 1 ATGAGGAGAGGCTTGAAGAGAAACCTGTATTGAAAGAGAGAGAGGTGGAG 60
Db 292 ATGATGGAGAGCTTGAAGAGAAACCTGTATTGAAAGAGAGAGAGGTGGAG 351
```

```
Qy 61 TGAAGTCCAGCAGGCTTCAAGCTACAGCAGAGCCCTGCTCCGATCTCTTAGCAAA 120
Db 352 TGAAGTCCAGCAGGCTTCAAGCTACAGCAGAGCCCTGCTCCGATCTCTTAGCAAA 411
```

```
Qy 121 GTCCAGCGAGGCTGGGCCCCGAGACCCAGCCTCCGAGGAGATGCTTCATTCATAC 180
Db 412 GTCCAGCGAGGCTGGGCCCCGAGACCCAGCCTCCGAGGAGATGCTTCATTCATAC 471
```

```
Qy 181 ACCCTCAACAGCTGAGGCTGACTTGGCTGGCTCCGCTCCGCTGCCCCCGAG 240
Db 472 ACCCTCAACAGCTGAGGCTGACTTGGCTGGCTCCGCTCCGCTGCCCCCGAG 531
```

```
Qy 241 CCCCTCTTCTGGGCGAGGAGATTTCTCCCTGTGAGCCACCATTTGCTCTACAG 300
Db 532 CCCCTCTTCTGGGCGAGGAGATTTCTCCCTGTGAGCCACCATTTGCTCTACAG 591
```

```
Qy 301 GAGCTGACACCTCCATGATGAGTGAAGCCCTCAGAAATCCAGTACTCCCTTGGC 360
Db 592 GAGCTGACACCTCCATGATGAGTGAAGCCCTCAGAAATCCAGTACTCCCTTGGC 651
```

```
Qy 361 CTCGAGATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 420
Db 652 CTCGAGATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 711
```

```
Qy 421 TACTTGGGGGACTCTGGCTGTGATGATCTTCTTGTGACATTGACATCTGCGTAGA 480
Db 712 TACTTGGGGGACTCTGGCTGTGATGATCTTCTTGTGACATTGACATCTGCGTAGA 771
```

```
Qy 481 AAGAGCCTGACCGGGCCCCCAGCAGGCTCTCAGAACCTTCTTGTGCGCCAGTTCT 540
Db 772 AAGAGCCTGACCGGGCCCCCAGCAGGCTCTCTCAACACTTCTGTGCGCCAGTTCT 831
```

```
Qy 541 TGGGAGTGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 591
Db 832 TGGGAGTGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 882
```

```
Qy 591 TGGGAGTGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 651
Db 882 TGGGAGTGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 941
```

```
Qy 651 TGGGAGTGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 711
Db 941 TGGGAGTGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 1000
```

```
Qy 711 TGGGAGTGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 771
Db 1000 TGGGAGTGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 1060
```

```
Qy 771 TGGGAGTGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 831
Db 1060 TGGGAGTGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 1120
```

```
Qy 831 TGGGAGTGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 891
Db 1120 TGGGAGTGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 1180
```

```
Qy 891 TGGGAGTGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 951
Db 1180 TGGGAGTGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 1240
```

```
Qy 951 TGGGAGTGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 1011
Db 1240 TGGGAGTGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 1300
```

```
Qy 1011 TGGGAGTGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 1071
Db 1300 TGGGAGTGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 1360
```

```
Qy 1071 TGGGAGTGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 1131
Db 1360 TGGGAGTGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 1420
```

```
Qy 1131 TGGGAGTGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 1191
Db 1420 TGGGAGTGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 1480
```

```
Qy 1191 TGGGAGTGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 1251
Db 1480 TGGGAGTGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 1540
```

```
Qy 1251 TGGGAGTGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 1311
Db 1540 TGGGAGTGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 1600
```

```
Qy 1311 TGGGAGTGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 1371
Db 1600 TGGGAGTGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 1660
```

```
Qy 1371 TGGGAGTGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 1431
Db 1660 TGGGAGTGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 1720
```

```
Qy 1431 TGGGAGTGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 1491
Db 1720 TGGGAGTGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 1780
```

```
Qy 1491 TGGGAGTGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 1551
Db 1780 TGGGAGTGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 1840
```

DE Human cDNA 5'-end sequence, SEQ ID NO: 121.
 XX Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
 XX Homo sapiens.
 XX EP1130094-A2.
 XX 05-SEP-2001.
 XX 07-JUL-2000; 2000BP-00114089.
 XX 08-JUL-1999; 99JP-00194486.
 PR 11-JAN-2000; 2000JP-00118774.
 PR 02-MAY-2000; 2000JP-00183765.
 XX (HELI-) HELIX RES INST.
 XX Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
 XX WPI; 2001-524255/58.
 XX 830 Primers useful for synthesizing full length cDNA clones and their use
 PT in genetic manipulation.
 XX Claim 2; SEQ ID NO 121; 1380bp + Sequence Listing; English.
 XX The invention relates to primers for synthesizing full length cDNA
 CC clones. 830 cDNA molecules encoding a human protein have been isolated
 CC and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
 CC been determined. Primers for synthesizing the full length cDNA are useful
 CC for clarifying the function of the protein encoded by the cDNA. The full
 CC length clones were obtained by construction of full length enriched cDNA
 CC libraries that were synthesised by the oligo-capping method. The primers
 CC enable the production of the full length cDNA easily without any special
 CC method. The present sequence is the nucleotide sequence of the 5'-end of
 CC a cDNA provided in the invention. Note: The sequence data for this patent
 CC did not form part of the printed specification, but was obtained in CD-
 CC ROM format directly from EPO
 XX
 XX Sequence 825 BP; 160 A; 247 C; 230 G; 183 T; 0 U; 5 Other;
 SQ
 Query Match 99.6%; Score 588.4; DB 4; Length 825;
 Best Local Similarity 99.7%; Pred. No. 2.2e-134;
 Matches 589; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 ATGAGGAGGAGCTTGAAGAGAAACACTGATTGGAAGAGAGAGAGAGGTGGAG 60
 DB 93 ATGTGTGGAGGCTTGAAGAGAAACACTGATTGGAAGAGAGAGAGAGGTGGAG 152
 QY 61 TGGAGTCCGAGAGGCTTGAAGAGAAACACTGATTGGAAGAGAGAGAGAGGTGGAG 120
 DB 153 TGGAGTCCGAGAGGCTTGAAGAGAAACACTGATTGGAAGAGAGAGAGAGGTGGAG 212
 QY 121 GTTCAGCGAGAGCTTGGAGGCTTGAAGAGAAACACTGATTGGAAGAGAGAGGTGGAG 180
 DB 213 GTTCAGCGAGAGCTTGGAGGCTTGAAGAGAAACACTGATTGGAAGAGAGAGGTGGAG 272
 QY 181 ACCCTTCAACAGAGCTTGAAGAGAAACACTGATTGGAAGAGAGAGAGGTGGAG 240
 DB 273 ACCCTTCAACAGAGCTTGAAGAGAAACACTGATTGGAAGAGAGAGAGGTGGAG 332
 QY 241 CCCCTTCTTGGGCGAGAGAGATTCTGCTGTGAGGACCAATGAGCTTATCCGAG 300
 DB 333 CCCCTTCTTGGGCGAGAGAGATTCTGCTGTGAGGACCAATGAGCTTATCCGAG 392
 QY 301 GAGCTGACACCTCCATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
 DB 393 GAGCTGACACCTCCATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 452
 QY 361 CTCGAGATGAAGTGCACCCAGAGCTGATTCAGTCTTTTGAAGAGCTTGAAGCTCCGAG 420

DB 453 CTCGAGATGAAGTGCACCCAGAGCTGATTCAGTCTTTTGAAGAGCTTGAAGCTCCGAG 512
 QY 421 TACTTGGGAGACTGTGGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
 DB 513 TACTTGGGAGACTGTGGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 572
 QY 481 AAGAGCTTGAAG 540
 DB 573 AAGAGCTTGAAG 632
 QY 541 TGGAGTGAATGAAGTGCATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 591
 DB 633 TGGAGTGAATGAAGTGCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 683
 RESULT 8
 AAK93617
 ID AAK93617 standard; cDNA; 825 BP.
 XX
 AC AAK93617;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human cDNA clone representative sequence, SEQ ID NO: 2077.
 XX
 XX Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
 XX Homo sapiens.
 XX EP1130094-A2.
 XX 05-SEP-2001.
 XX 07-JUL-2000; 2000BP-00114089.
 XX 08-JUL-1999; 99JP-00194486.
 PR 11-JAN-2000; 2000JP-00118774.
 PR 02-MAY-2000; 2000JP-00183765.
 XX (HELI-) HELIX RES INST.
 XX Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
 XX WPI; 2001-524255/58.
 XX 830 Primers useful for synthesizing full length cDNA clones and their use
 PT in genetic manipulation.
 XX Example 11; SEQ ID NO 2077; 1380bp + Sequence Listing; English.
 XX The invention relates to primers for synthesizing full length cDNA
 CC clones. 830 cDNA molecules encoding a human protein have been isolated
 CC and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
 CC been determined. Primers for synthesizing the full length cDNA are useful
 CC for clarifying the function of the protein encoded by the cDNA. The full
 CC length clones were obtained by construction of full length enriched cDNA
 CC libraries that were synthesised by the oligo-capping method. The primers
 CC enable the production of the full length cDNA easily without any special
 CC method. The present sequence was used as the representative sequence
 CC from a human clone which was used in homology searches to identify the
 CC clone. Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in CD-ROM format directly from
 CC EPO
 XX
 XX Sequence 825 BP; 160 A; 247 C; 230 G; 183 T; 0 U; 5 Other;
 SQ
 Query Match 99.6%; Score 588.4; DB 4; Length 825;
 Best Local Similarity 99.7%; Pred. No. 2.2e-134;
 Matches 589; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 ATGAGGAGGAGCTTGAAGAGAAACACTGATTGGAAGAGAGAGAGAGGTGGAG 60

```
Db 93 ATGTGGAGGCTTGAAGAGAAACACTGTATTTGGAAGAGAGAGAGAGTGGAG 152
Oy 61 TGGAGTCCAGCAGGCTTTCAGAGCTACAGAGAGCCCTGCTCGCATCTCCCTAGACAA 120
Db 153 TGGAGTCCAGCAGGCTTTCAGAGCTACAGAGAGCCCTGCTCGCATCTCCCTAGACAA 212
Oy 121 GTCCAGCCGAGCTTGGGCCCCCGAGACCCAGCTTCGCGAGAGATGCTCATTCATAC 180
Db 213 GTCCAGCCGAGCTTGGGCCCCCGAGACCCAGCTTCGCGAGAGATGCTCATTCATAC 272
Oy 181 ACCCTCCAAAGCTGCAAGCTGCACTGCGCTGCGCTCCGCGCTCCGCTCCGCGCGAG 240
Db 273 ACCCTCCAAAGCTGCAAGCTGCACTGCGCTGCGCTCCGCGCTCCGCGCGAG 332
Oy 241 CCCCTCTTCTGGGCGAGAGATTTCTCTGTCAGCAGCAGCATTTGCTCTATCTCAG 300
Db 333 CCCCTCTTCTGGGCGAGAGATTTCTCTGTCAGCAGCAGCATTTGCTCTATCTCAG 392
Oy 301 GAGCTGACACCTTCATGATGAGAGCTGAGCCCTCAGAAATCCAGTACTCCCTTGGC 360
Db 393 GAGCTGACACCTTCATGATGAGAGCTGAGCCCTCAGAAATCCAGTACTCCCTTGGC 452
Oy 361 CTCGAGATGAGATGAGAGCTGAGAGCTGATCCAGCTTTCTTGAAGCTTGAAGCTCCCG 420
Db 453 CTCGAGATGAGATGAGAGCTGAGAGCTGATCCAGCTTTCTTGAAGCTTGAAGCTCCCG 512
Oy 421 TACTTGGGGGAACTCTGAGCTGATGATCTTTCTTGAACATTTGACATCTCGGTAGA 480
Db 513 TACTTGGGGGAACTCTGAGCTGATGATCTTTCTTGAACATTTGACATCTCGGTAGA 572
Oy 481 AAGAGCCTGCAAGGAGCCGACCAAGAGCTCTCAAACTCTTCTGTCGCGCAAGTTCT 540
Db 573 AAGAGCCTGCAAGGAGCCGACCAAGAGCTCTCAAACTCTTCTGTCGCGCAAGTTCT 632
Oy 541 TGGGAGTGAATGAATGATGATCATCATGAAATCATTTGGGGTCTTAA 591
Db 633 TGGGAGTGAATGAATGATGATCATCATGAAATCATTTGGGGTCTTAA 683

RESULT 9
ADL30044
ID ADL30044 standard; cDNA; 825 BP.
AC XX
ADL30044;
AC XX
DT 20-MAY-2004 (first entry)
XX XX
DE 3' end of a representative human cDNA cluster SegID 2077.
XX XX
KW human; medicine; signal transduction; glycoprotein; transcription;
KW oligo-capping method; ss.
XX
OS Homo sapiens.
XX
PN EP1396543-A2.
XX
PD 10-MAR-2004.
XX
PF 07-JUL-2000; 2003BP-00025638.
XX
PR 08-JUL-1999; 99JP-00194486.
PR 11-JAN-2000; 2000JP-0018774.
PR 02-MAY-2000; 2000JP-00183865.
PR 07-JUL-2000; 2000BP-00114089.
XX
PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX WPI; 2004-204755/20.
XX
PT New oligonucleotide primers (830 cDNAs) useful for synthesizing full
```

```
PT length human cDNAs.
XX
PS Example 18; SEQ ID NO 2077; 1340bp; English.
XX
CC This invention relates to a novel primers useful for synthesizing full
CC length cDNA molecules that encode human proteins. Specifically, it refers
CC to secretory or membrane proteins that are potential therapeutic agents/
CC target molecules in the field of medicine, and in particular genes
CC encoding proteins that are associated with signal transduction,
CC glycoproteins and transcription. The present invention describes a method
CC for efficiently cloning a full length human cDNA from both the 5' and 3'
CC ends using the oligo-capping method. This polynucleotide sequence is the
CC 3' end of a representative human DNA cluster of the invention.
XX
XX Sequence 825 BP; 160 A; 247 C; 230 G; 183 T; 0 U; 5 Other;
Query Match 99.6%; Score 588.4; DB 12; Length 825;
Best Local Similarity 99.7%; Pred. No. 2.2e-134;
Matches 589; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Oy 1 ATGAGGAGGAGCTTGAAGAGAAACACTGTATTTGGAAGAGAGAGAGAGTGGAG 60
Db 93 ATGTGGAGGCTTGAAGAGAAACACTGTATTTGGAAGAGAGAGAGAGTGGAG 152
Oy 61 TGGAGTCCAGCAGGCTTTCAGAGCTACAGAGAGCCCTGCTCGCATCTCCCTAGACAA 120
Db 153 TGGAGTCCAGCAGGCTTTCAGAGCTACAGAGAGCCCTGCTCGCATCTCCCTAGACAA 212
Oy 121 GTCCAGCCGAGCTTGGGCCCCCGAGACCCAGCTTCGCGAGAGATGCTCATTCATAC 180
Db 213 GTCCAGCCGAGCTTGGGCCCCCGAGACCCAGCTTCGCGAGAGATGCTCATTCATAC 272
Oy 181 ACCCTCCAAAGCTGCAAGCTGCACTGCGCTGCGCTCCGCGCTCCGCTCCGCGCGAG 240
Db 273 ACCCTCCAAAGCTGCAAGCTGCACTGCGCTGCGCTCCGCGCTCCGCGCGAG 332
Oy 241 CCCCTCTTCTGGGCGAGAGATTTCTCTGTCAGCAGCAGCATTTGCTCTATCTCAG 300
Db 333 CCCCTCTTCTGGGCGAGAGATTTCTCTGTCAGCAGCAGCATTTGCTCTATCTCAG 392
Oy 301 GAGCTGACACCTTCATGATGAGAGCTGAGCCCTCAGAAATCCAGTACTCCCTTGGC 360
Db 393 GAGCTGACACCTTCATGATGAGAGCTGAGCCCTCAGAAATCCAGTACTCCCTTGGC 452
Oy 361 CTCGAGATGAGATGAGAGCTGAGAGCTGATCCAGCTTTCTTGAAGCTTGAAGCTCCCG 420
Db 453 CTCGAGATGAGATGAGAGCTGAGAGCTGATCCAGCTTTCTTGAAGCTTGAAGCTCCCG 512
Oy 421 TACTTGGGGGAACTCTGAGCTGATGATCTTTCTTGAACATTTGACATCTCGGTAGA 480
Db 513 TACTTGGGGGAACTCTGAGCTGATGATCTTTCTTGAACATTTGACATCTCGGTAGA 572
Oy 481 AAGAGCCTGCAAGGAGCCGACCAAGAGCTCTCAAACTCTTCTGTCGCGCAAGTTCT 540
Db 573 AAGAGCCTGCAAGGAGCCGACCAAGAGCTCTCAAACTCTTCTGTCGCGCAAGTTCT 632
Oy 541 TGGGAGTGAATGAATGATGATCATCATGAAATCATTTGGGGTCTTAA 591
Db 633 TGGGAGTGAATGAATGATGATCATCATGAAATCATTTGGGGTCTTAA 683

RESULT 10
ADL28088
ID ADL28088 standard; cDNA; 825 BP.
AC XX
ADL28088;
AC XX
DT 20-MAY-2004 (first entry)
XX XX
DE 5' end of a human cDNA molecule SegID 121.
XX XX
KW human; medicine; signal transduction; glycoprotein; transcription;
KW oligo-capping method; ss.
XX
```

XX Homo sapiens.
 OS
 XX
 XX EP1396543-A2.
 XX
 XX
 PD 10-MAR-2004.
 XX
 XX
 PP 07-JUL-2000; 2003EP-00025638.
 XX
 XX 08-JUL-1999; 99JP-00194486.
 PR 11-JAN-2000; 2000JP-00118774.
 PR 02-MAY-2000; 2000JP-00183865.
 PR 07-JUL-2000; 2000EP-00114089.
 XX
 XX (REAS-) RES ASSOC BIOTECHNOLOGY.
 XX
 PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
 PI Makamatsu A, Sugiyama T, Negai K, Kojima S, Otsuki T, Koga H;
 XX
 XX WPI, 2004-204755/20.
 XX
 PT New oligonucleotide primers (830 CDNAs) useful for synthesizing full
 PT length human CDNAs.
 PS
 PS Disclosure; SEQ ID NO 121, 1340pp; English.
 XX
 XX This invention relates to a novel primers useful for synthesizing full
 CC length cDNA molecules that encode human proteins. Specifically, it refers
 CC to secretory or membrane proteins that are potential therapeutic agents/
 CC target molecules in the field of medicine, and in particular genes
 CC encoding proteins that are associated with signal transduction,
 CC glycoproteins and transcription. The present invention describes a method
 CC for efficiently cloning a full length human cDNA from both the 5' and 3'
 CC ends using the oligo-capping method. This polynucleotide sequence is the
 CC 5' end of a full length human cDNA sequence of the invention.
 XX
 SQ Sequence 825 BP, 160 A, 247 C, 230 G, 183 T, 0 U, 5 Other;
 Query Match 99.6%; Score 588.4; DB 12; Length 825;
 Best Local Similarity 99.7%; Pred. No. 2.2e-134;
 Matches 589; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAGGAGAGGCTTGAAGAGAAACCTGTGATTTGAAAGAGAGAGAGAGGTGGAG 60
 DB 93 ATGAGGAGAGGCTTGAAGAGAAACCTGTGATTTGAAAGAGAGAGAGAGGTGGAG 152
 QY 61 TGGAGTCCAGCAGGCGCTTCAAGGCTACAGAGAAAGCCCTGCTCCGATCTCCCTAGACAA 120
 DB 153 TGGAGTCCAGCAGGCGCTTCAAGGCTACAGAGAAAGCCCTGCTCCGATCTCCCTAGACAA 212
 QY 121 GTCCAGGCGAGGCTGGGCCCCCGAGAGACCCAGAGCTCCGAGGATGTCTTCATCCATAC 180
 DB 213 GTCCAGGCGAGGCTGGGCCCCCGAGAGACCCAGAGCTCCGAGGATGTCTTCATCCATAC 272
 QY 181 ACCCTTCAACAGCTGCAAGGCTGACCTTGGCTGCGCTCCGAGCCCTGCTGCTCCGAG 240
 DB 273 ACCCTTCAACAGCTGCAAGGCTGACCTTGGCTGCGCTCCGAGCCCTGCTGCTCCGAG 332
 QY 241 CCCCTTCTTCTGGGAGAGAGATTTTCCCTGTAGAGACCAATTTGGCTTATCTCCAGG 300
 DB 333 CCCCTTCTTCTGGGAGAGAGATTTTCCCTGTAGAGACCAATTTGGCTTATCTCCAGG 392
 QY 301 GAGCTGACACCTTCATGTAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
 DB 333 GAGCTGACACCTTCATGTAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 452
 QY 351 CTCGAGATGAAGTGCACCCAGAGCTGATTCATTCCTTGAAGAGAGAGAGAGAGAGAG 420
 DB 453 CTCGAGATGAAGTGCACCCAGAGCTGATTCATTCCTTGAAGAGAGAGAGAGAGAGAG 512
 QY 421 TACTTGGGGAGAGCTTGGGCTGTGATGATCTTTCTTGAAGATGACATCTGCGGTAGAA 480
 DB 513 TACTTGGGGAGAGCTTGGGCTGTGATGATCTTTCTTGAAGATGACATCTGCGGTAGAA 572

QY 481 AAGAGCTTGACAGCGGCCCCCAGAGAGCTTCTTCAACAACTTTCTGTGCCCCAGTTCT 540
 DB 573 AAGAGCTTGACAGCGGCCCCCAGAGAGCTTCTTCAACAACTTTCTGTGCCCCAGTTCT 632
 QY 541 TGGGAGTGAATGAATGATGATCATCATCATGAAATCATTTCTGGGCTCTAA 591
 DB 633 TGGGAGTGAATGAATGATGATCATCATCATGAAATCATTTCTGGGCTCTAA 683

RESULT 11
 ACH91703
 ID ACH91703 standard; DNA; 598 BP.
 XX
 XX ACH91703;
 XX
 XX 29-JUL-2004 (first entry)
 XX
 XX Human genome derived single exon probe #24898.
 XX
 XX Human; probe; sa; gene expression; single exon probe; microarray;
 KW alternative splicing event; genomic alteration.
 XX
 XX Homo sapiens.
 OS
 XX US2003194704-A1.
 XX
 XX 16-OCT-2003.
 XX
 XX 03-APR-2002; 2002US-00029386.
 XX
 XX 03-APR-2002; 2002US-00029386.
 XX
 XX (PENN/) PENN S G.
 PA (RANK/) RANK D R.
 PA (HANZ/) HANZEL D K.
 XX
 PI Penn SG, Rank DR, Hanzel DK;
 PI WPI, 2004-119264/12.
 DR
 XX
 PT New human genome-derived single exon nucleic acid probes useful for human
 PT gene expression analysis, for identifying or characterizing alternative
 PT splicing events, for assessing genomic alterations or as tools for
 PT surveying tissues.
 XX
 XX
 PS Claim 1; SEQ ID NO 24898; 80pp; English.

CC The invention relates to a nucleic acid probe for measuring human gene
 CC expression, comprising any of the 27,400 fully defined nucleotide
 CC sequences in the specification, or their complements or fragments, and
 CC encoding at least 8 amino acids of any of the 688 amino acid sequences
 CC fully defined in the specification. The probe is a single exon probe that
 CC hybridizes under high stringency conditions to a nucleic acid molecule
 CC expressed in human cells or tissues. Also included are a spatially-
 CC addressable set of single exon nucleic acid probes for measuring human
 CC gene expression (comprising a plurality of single exon nucleic acid
 CC probes cited above, where each of the plurality of probes is separately
 CC and addressably isolatable or amplifiable from the plurality), a single
 CC exon microarray for measuring human gene expression, a method of
 CC measuring human gene expression, a vector comprising the single exon
 CC probe cited above, an ORF-encoded peptide comprising at least 8
 CC contiguous amino acids of any of the above-mentioned amino acid
 CC sequences (optionally with conservative amino acid substitutions), an
 CC isolated antibody that binds specifically to a peptide cited above,
 CC methods of selling and/or licensing single exon probes or microarrays to
 CC a customer desiring to measure gene expression, a method of providing
 CC human gene expression data by subsequence, and a computer-readable
 CC storage medium which contains a database having a plurality of records
 CC (each record including data on the expression of a single exon probe
 CC cited above. The probe, methods and apparatus are useful in gene
 CC expression analysis. The probes may be used as tools for surveying
 CC tissues to detect the presence of expressed messages that contain their

CC specific exon, or in constructing genome-derived single exon microarrays.
CC In addition, the probes are used in identifying and characterizing
CC alternative splicing events, in detecting and characterizing gross
CC alterations in the genomic locus that includes their exon, in assessing
CC smaller genomic alterations, in priming the synthesis of nucleic acids,
CC or in expressing the ORF-encoded peptide. The present sequence is a human
CC single exon probe of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from USPTO at
CC segdata.uspto.gov/sequence.html?docID=20030194704
XX

Seq Sequence 598 BP; 123 A; 197 C; 158 G; 120 T; 0 U; 0 Other;

Query Match 97.9%; Score 578.4; DB 12; Length 598;

Best Local Similarity 99.7%; Pred. No. 5.9e-132;

Matches 590; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 1 ATGAGGAGAGGCTTGAAGAAACACTGTGATTTGGAAGAGAGAGAGAGGTGGAG 60
DB 7 ATGTGGGAGGCTTGAAGAAACACTGTGATTTGGAAGAGAGAGAGAGGTGGAG 66
QY 61 TGGAGTCCAGAGGCTTCAAGAGCTACAGCAAGGCTCGCATCTCCCTAGACAA 120
DB 67 TGAAGTCCAGAGGCTTCAAGAGCTACAGCAAGGCTCGCATCTCCCTAGACAA 126
QY 121 GTTCAGCGAGGCTTGGGCGCCCGAGAGCCAGCCTCCGAGGAGATGCTTCATCAAC 180
DB 127 GTTCAGCGAGGCTTGGGCGCCCGAGAGCCAGCCTCCGAGGAGATGCTTCATCAAC 186
QY 181 ACCCTTCAACAGCTGAGAGGCTGCACTTGGCTGGCTCCGCTCCGCTGCTGCTG 239
DB 187 ACCCTTCAACAGCTGAGAGGCTGCACTTGGCTGGCTCCGCTCCGCTGCTGCTG 246
QY 240 GCCCTCTTCCGCGGAGAGAGATTTCCCTGTCAGCAGCAGCATTTGCTATCTGAG 299
DB 247 GCCCTCTTCCGCGGAGAGAGATTTCCCTGTCAGCAGCAGCATTTGCTATCTGAG 306
QY 300 GGAGCTGAGCACTTCATGATGAGAGCTGAGGCGCCCTCGAATCCAGTATCCCTTGG 359
DB 307 GGAGCTGAGCACTTCATGATGAGAGCTGAGGCGCCCTCGAATCCAGTATCCCTTGG 366
QY 360 CCTCCAGATGAAGTGCACCCAGCAGCTGATCAGTCTTCTTGAAGCTCTGAGCTCCG 419
DB 367 CCTCCAGATGAAGTGCACCCAGCAGCTGATCAGTCTTCTTGAAGCTCTGAGCTCCG 426
QY 420 GTACTTTGGGGAGCTCTGGCTTGAATGACTTTCTTGAATGATGATCTGCGGTAGA 479
DB 427 GTACTTTGGGGAGCTCTGGCTTGAATGACTTTCTTGAATGATGATCTGCGGTAGA 486
QY 480 AAAAGAGCTGAGCGGGCCCGAGAGGCTCTCAACACCTCTGTCGCGCCAGGTTG 539
DB 487 AAAAGAGCTGAGCGGGCCCGAGAGGCTCTCTCAACACCTCTGTCGCGCCAGGTTG 546
QY 540 TTGGAGTGAATGAATGATGATCAATCAATGAAATCAATCTTGGGGTCTTAA 591
DB 547 TTGGAGTGAATGAATGATGATCAATCAATGAAATCAATCTTGGGGTCTTAA 598

RESULT 12
AAH34147
ID AAH34147 standard; cDNA; 1596 BP.

XX AC AAH34147;

DT 03-SEP-2001 (first entry)

DE Human colon cancer antigen encoding cDNA SEQ ID NO:1229.

XX Human; colon cancer; colon cancer antigen; diagnosis; detection;

KM colorectal carcinoma; ss.

XX Homo sapiens.

OS

XX

PN WO200122920-A2.
XX
XX 05-APR-2001.
XX
XX 28-SEP-2000; 2000WO-US026524.
XX
XX 29-SEP-1999; 99US-0157137P.
PR 03-NOV-1999; 99US-0163280P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Ruben SM, Barash SC, Birse CE, Rosen CA;
PI WPI; 2001-235357/24.
DR P-PSDB; AAG74742.
PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT useful for preventing, diagnosing and/or treating colorectal cancers.
XX
XX Claim 1; Page 3003-3004; 9803pp; English.

AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
cancer-associated nucleic acid molecules (N) and proteins (P), where the
cancer proteins are collectively known as colon cancer antigens. The colon
cancer antigens have cytoskeletal activity and can be used in gene therapy
CC and vaccine production. N and P may be used in the prevention, diagnosis
CC and treatment of diseases associated with inappropriate P expression. For
CC example, N and P may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of P by expressing inactive proteins or to
CC supplement the patient's own production of P. Additionally, N may be used
CC to produce the colon cancer-associated P, by inserting the nucleic acids
CC into a host cell and culturing the cell to express the proteins. N and P
CC can be used in the prevention, diagnosis and treatment of colorectal
cancer and cancers. AAH37196 to AAH37204 and AAB77789 represent
cancer sequences used in the exemplification of the present invention. N.B.
CC Pages 666 to 682 and page 7053 of the sequence listing were missing at
CC time of publication, meaning no sequences are present for SEQ ID NO:1027
CC to 1052, 7921 and 7922
XX

Seq Sequence 1596 BP; 354 A; 418 C; 410 G; 409 T; 0 U; 5 Other;

Query Match 97.5%; Score 576.2; DB 4; Length 1596;

Best Local Similarity 99.3%; Pred. No. 2.6e-131;

Matches 587; Conservative 2; Mismatches 0; Indels 2; Gaps 1;

QY 1 ATGAGGAGAGGCTTGAAGAAACACTGTGATTTGGAAGAGAGAGAGGTGGAG 60
DB 345 ATGAGGAGAGGCTTGAAGAAACACTGTGATTTGGAAGAGAGAGAGGTGGAG 404
QY 61 TGGAGTCCAGAGGCTTCAAGAGCTACAGCAAGGCTCGCATCTCCCTAGACAA 120
DB 67 TGAAGTCCAGAGGCTTCAAGAGCTACAGCAAGGCTCGCATCTCCCTAGACAA 126
QY 121 GTTCAGCGAGGCTTGGGCGCCCGAGAGCCAGCCTCCGAGGAGATGCTTCATCAAC 180
DB 127 GTTCAGCGAGGCTTGGGCGCCCGAGAGCCAGCCTCCGAGGAGATGCTTCATCAAC 186
QY 181 ACCCTTCAACAGCTGAGAGGCTGCACTTGGCTGGCTCCGCTCCGCTGCTGCTG 239
DB 187 ACCCTTCAACAGCTGAGAGGCTGCACTTGGCTGGCTCCGCTCCGCTGCTGCTG 246
QY 240 GCCCTCTTCCGCGGAGAGAGATTTCCCTGTCAGCAGCAGCATTTGCTATCTGAG 299
DB 247 GCCCTCTTCCGCGGAGAGAGATTTCCCTGTCAGCAGCAGCATTTGCTATCTGAG 306
QY 300 GGAGCTGAGCACTTCATGATGAGAGCTGAGGCGCCCTCGAATCCAGTATCCCTTGG 359
DB 307 GGAGCTGAGCACTTCATGATGAGAGCTGAGGCGCCCTCGAATCCAGTATCCCTTGG 366
QY 360 CCTCCAGATGAAGTGCACCCAGCAGCTGATCAGTCTTCTTGAAGCTCTGAGCTCCG 419
DB 367 CCTCCAGATGAAGTGCACCCAGCAGCTGATCAGTCTTCTTGAAGCTCTGAGCTCCG 426
QY 420 GTACTTTGGGGAGCTCTGGCTTGAATGACTTTCTTGAATGATGATCTGCGGTAGA 479
DB 427 GTACTTTGGGGAGCTCTGGCTTGAATGACTTTCTTGAATGATGATCTGCGGTAGA 486
QY 480 AAAAGAGCTGAGCGGGCCCGAGAGGCTCTCAACACCTCTGTCGCGCCAGGTTG 539
DB 487 AAAAGAGCTGAGCGGGCCCGAGAGGCTCTCTCAACACCTCTGTCGCGCCAGGTTG 546
QY 540 TTGGAGTGAATGAATGATGATCAATCAATGAAATCAATCTTGGGGTCTTAA 591
DB 547 TTGGAGTGAATGAATGATGATCAATCAATGAAATCAATCTTGGGGTCTTAA 598

QY 421 TACTTGGGGGACTCTGGCCCTGGATGATCTTTCTTGAACATTGCGGTAA 480
DB 763 TACTTGGGGGACTCTGGCCCTGGATGATCTTTCTTGAACATTGCGGTAA 822
QY 481 AAGAGGCTGACCGGGCCCGACAGAGCTCTCTCAACCTCTTGTGGCCGAGTTCT 540
DB 823 AAGAGGCTGACCGGGCCCGACAGAGCTCTCTCAACCTCTTGTGGCCGAGTTCT 882
QY 541 TGGAGTGAATGAATGATGATCATCATGAAATCATTTGGGTCTAA 591
DB 883 TGGAGTGAATGAATGATGATCATCATGAAATCATTTGGGTCTAA 933

RESULT 13
ABL83086
ID ABL83086 standard; cDNA, 506 BP.
AC ABL83086;
DT 17-MAY-2002 (first entry)
XX
DE Human ovarian cancer related cDNA clone SEQ ID NO:6064.
XX
KM Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.
XX
OS Homo sapiens.
XX
PN MO200192581-A2.
XX
PD 06-DEC-2001.
XX
PP 29-MAY-2001; 2001MO-US017756.
XX
PR 26-MAY-2000; 2000US-0207484P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Algate PA, Harlocker SL, Jones R;
XX
DR WPI, 2002-122075/16.
XX
PT Composition for therapy and diagnosis of ovarian cancer comprising
PT polypeptide of a ovarian tumor polypeptide, polynucleotide encoding
PT polypeptide, antibody specific to polypeptide or T cell expressing
PT polypeptide.
PS Claim 1; SEQ ID NO 6064; 489pp; English.
XX
XX The present invention describes a composition (I) comprising: carriers
CC and immunostimulants; and a polypeptide (II) of a ovarian tumour
CC polypeptide encoded by a polynucleotide (III) having a cDNA sequence (S1)
CC from the 10912 nucleotide sequences as given in ABL77023 to ABL87934,
CC (III) encoding (II) having a sequence (S2), a T cell population of (II),
CC or antigen presenting cells that express (II). (I) has cytostatic
CC activity. An oligonucleotide (IV) that hybridises to (S1) can be used for
CC detecting ovarian cancer in a patient's biological sample preferably
CC serum or ovarian tissue. The method comprises contacting a biological
CC sample from a patient with (IV), detecting the amount of polynucleotide
CC hybridising to (IV) and comparing the amount to a predetermined cutoff
CC value and thereby detecting ovarian cancer in the patient, where the
CC amount of polynucleotide hybridising to (IV) is detected preferably by
CC polymerase chain reaction (PCR). (I) comprising (III) and/or (II) is
CC useful for stimulating and/or expanding T cells specific for an ovarian
CC tumour protein comprising contacting T cells with (III) or (II). (III) is
CC useful in design and preparation of ribozyme molecules for inhibiting
CC expression of the tumour polypeptide and proteins in tumour cells; and
CC to isolate a full length gene from a suitable library e.g., a tumour cDNA
CC library using well known techniques
SQ Sequence 506 BP; 98 A; 184 C; 124 G; 100 T; 0 U; 0 Other;

Query Match 80.0%; Score 473; DB 6; Length 506;

Best Local Similarity 98.6%; Pred. No. 4,1e-106;
Matches 488; Conservative 0; Mismatches 5; Indels 2; Gaps 1;
QY 60 GTGAGTCACAGAGCCCTTCAAGACTACACAGAGCCCTGCTCGATCTCCAGACAA 119
DB 14 GGGAGTCACAGAGCCCTTCAAGACTACACAGAGCCCTGCTCGATCTCCAGACAA 73
QY 120 AGTCAGGCGAGCCTGGGCCCCCGACAGACCCAGCTCTCGAGGACATGTCCTCATCA 179
DB 74 AGTCAGGCGC--CCTGGGCCCCCGACAGACCCAGCTCTCGAGGACATGTCCTCATCA 131
QY 180 CACCTCCACAGAGCTGACAGCTGACCTTGCTGCTCCGCCCCCTGCTGCCCCCGA 239
DB 132 CACCTCCACAGAGCTGACAGATGACCTTGCTGCTCCGCCCCCTGCTGCCCCCGA 191
QY 240 GCCCTCTTCTGGGCGAGAGATTTCTCCCTGTACACCAACATTTGCTTATCTGAG 299
DB 192 GCCCTCTTCTGGGCGAGAGATTTCTCCCTGTACACCAACATTTGCTTATCTGAG 251
QY 300 GAGGCTGACACCTTCATGATGATGGAGCTGAGCCCCCTCAAGATCCAGTACTCCCTTGG 359
DB 252 GAGGCTGACACCTTCATGATGATGGAGCTGAGCCCCCTCAAGATCCAGTACTCCCTTGG 311
QY 360 CTTCCAGAAATGAATGACCCAGCCAGCTGATTCAGTCTTTTGAAGAGCTGAGCTCCG 419
DB 312 CTTCCAGAAATGAATGACCCAGCCAGCTGATTCAGTCTTTTGAAGAGCTGAGCTCCG 371
QY 420 GTACTTGGGGGACTCTGAGCTGATGATGATCTTTCTTGAACATTGACATCTGCGGTAA 479
DB 372 GTACTTGGGGGACTCTGAGCTGATGATGATCTTTCTTGAACATTGACATCTGCGGTAA 431
QY 480 AAGAGGCTGACCGGGCCCGACAGAGCTCTCTCAACCTCTTGTGGCCGAGTTCT 539
DB 432 AAGAGGCTGACCGGGCCCGACAGAGCTCTCTCAACCTCTTGTGGCCGAGTTCT 491
QY 540 TTGGAGTGAATGA 554
DB 492 TTGGAGTGAATGA 506

RESULT 14
ABL83261
ID ABL83261 standard; cDNA, 205 BP.
AC ABL83261;
DT 17-MAY-2002 (first entry)
XX
DE Human ovarian cancer related cDNA clone SEQ ID NO:6239.
XX
KM Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200192581-A2.
XX
PD 06-DEC-2001.
XX
PP 29-MAY-2001; 2001MO-US017756.
XX
PR 26-MAY-2000; 2000US-0207484P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Algate PA, Harlocker SL, Jones R;
XX
DR WPI, 2002-122075/16.
XX
PT Composition for therapy and diagnosis of ovarian cancer comprising
PT polypeptide of a ovarian tumor polypeptide, polynucleotide encoding
PT polypeptide, antibody specific to polypeptide or T cell expressing
PT polypeptide.
XX

PS Claim 1; SEQ ID NO 6239; 489bp; English.
XX
CC The present invention describes a composition (I) comprising: carriers
CC and immunostimulants; and a polypeptide (II) of an ovarian tumour
CC polypeptide encoded by a polynucleotide (III) having a cDNA sequence (S1)
CC from the 10912 nucleotide sequence as given in AB177023 to AB187934,
CC (III) encoding (II) having a sequence (S2), a T cell population of (II),
CC or antigen presenting cells that express (II). (I) has cytostatic
CC activity. An oligonucleotide (IV) that hybridizes to (S1) can be used for
CC detecting ovarian cancer in a patient's biological sample preferably
CC serum or ovarian tissue. The method comprises contacting a biological
CC sample from a patient with (IV), detecting the amount of polynucleotide
CC hybridizing to (IV) and comparing the amount to a predetermined cutoff
CC value and thereby detecting ovarian cancer in the patient, where the
CC amount of polynucleotide hybridizing to (IV) is detected preferably by
CC polymerase chain reaction (PCR). (I) comprising (III) and/or (II) is
CC useful for stimulating and/or expanding T cells specific for an ovarian
CC tumour protein comprising contacting T cells with (III) or (II). (II) is
CC useful in design and preparation of ribozyme molecules for inhibiting
CC expression of the tumour polypeptides and proteins in tumour cells; and
CC to isolate a full length gene from a suitable library e.g., a tumour cDNA
CC library using well known techniques
XX
SQ Sequence 205 BP; 36 A; 89 C; 45 G; 35 T; 0 U; 0 Other;
Query Match 31.6%; Score 186.6; DB 6; Length 205;
Best Local Similarity 97.1%; Pred. No. 5,9e-36;
Matches 201; Conservative 0; Mismatches 4; Indels 2; Gaps 1;
QY 63 GAGTCAGAGGCTTCAGAGCTACAGAGAGCCCTGCTCCGATCTCCCTAGACAAAGT 122
DB 1 GAGTCAGAGGCTTCAGAGCTACAGAGAGCCCTGCTCCGATCTCCCTAGACAAAGT 60
QY 123 CGAGGCGAGCCCTGGGCCCCCGAGCACCCAGCCTCGGACGATGCTCATCAATACAC 182
DB 61 CGAGGCGC--CTGGGCCCCCGAGCACCCAGCCTCGGACGATGCTCATCAATACAC 118
QY 183 CCTCCAAAGCTGAGAGCTGCACTTGGCTCCGCGCCCTGACCTGCCCCCGAGGCC 242
DB 119 CCTCCAAAGCTGAGAGCTGCACTTGGCTCCGCGCCCTGACCTGCCCCCGAGGCC 178
QY 243 CCTCTTCTCTGGGCGAGAGGAGATTTCTC 269
DB 179 CCTCTTCTCTGGGCGAGAGGAGATTTCTC 205
RESULT 15
ACH77983
ID ACH77983 standard; DNA; 526 BP.
XX
AC ACH77983;
XX
DT 29-JUL-2004 (first entry)
XX
DE Human genome derived single exon probe #11178.
XX
KW Human; probe; ss; gene expression; single exon probe; microarray;
XX alternative splicing event; genomic alteration.
XX
OS Homo sapiens.
XX
PN US2003194704-A1.
PD 16-OCT-2003.
XX
PF 03-APR-2002; 2002US-00029386.
XX
PR 03-APR-2002; 2002US-00029386.
XX
PA (PENN/) PENN S G.
PA (RANK/) RANK D R.
PA (HANZ/) HANZEL D K.
XX

PI Penn SG, Rank DR, Hanzel DK;
DR WPI; 2004-119264/12.
XX
PT New human genome-derived single exon nucleic acid probes useful for human
PT gene expression analysis, for identifying or characterizing alternative
PT splicing events, for assessing genomic alterations or as tools for
PT surveying tissues.
XX
PS Claim 15; SEQ ID NO 11178; 80bp; English.
XX
CC The invention relates to a nucleic acid probe for measuring human gene
CC expression, comprising any of the 27,400 fully defined nucleotide
CC sequences in the specification, or their complements or fragments, and
CC encoding at least 8 amino acids of any of the 688 amino acid sequences
CC fully defined in the specification. The probe is a single exon probe that
CC hybridizes under high stringency conditions to a nucleic acid molecule
CC expressed in human cells or tissues. Also included are a spatially-
CC addressable set of single exon nucleic acid probes for measuring human
CC gene expression (comprising a plurality of single exon nucleic acid
CC probes cited above, where each of the plurality of probes is separately
CC and addressably isolatable or amplifiable from the plurality), a single
CC exon microarray for measuring human gene expression, a method of
CC measuring human gene expression, a vector comprising the single exon
CC probe cited above, an ORF-encoded peptide comprising at least 8
CC contiguous amino acids of any of the above-mentioned amino acid
CC sequences (optionally with conservative amino acid substitutions), an
CC isolated antibody that binds specifically to a peptide cited above,
CC methods of selling and/or licensing single exon probes or microarrays to
CC a customer desiring to measure gene expression, a method of providing
CC human gene expression data by subscription, and a computer-readable
CC storage medium which contains a database having a plurality of records
CC (each record including data on the expression of a single exon probe
CC cited above. The probe, methods and apparatus are useful in gene
CC expression analysis. The probes may be used as tools for surveying
CC tissues to detect the presence of expressed messages that contain their
CC specific exon, or in constructing genome-derived single exon microarrays.
CC In addition, the probes are used in identifying and characterizing
CC alternative splicing events, in detecting and characterizing gross
CC alterations in the genomic locus that includes their exon, in assessing
CC smaller genomic alterations, in priming the synthesis of nucleic acids,
CC or in expressing the ORF-encoded peptide. The present sequence is a human
CC single exon probe of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030194704
XX
SQ Sequence 526 BP; 105 A; 133 C; 134 G; 154 T; 0 U; 0 Other;
Query Match 28.1%; Score 166; DB 12; Length 526;
Best Local Similarity 100.0%; Pred. No. 8.3e-31;
Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 426 GGGGAGCTCTGGCTCGATGATCTTCTTCTGACATTTGACATCTGCGGTGAAGAAGA 485
DB 1 GGGGAGCTCTGGCTCGATGATCTTCTTCTGACATTTGACATCTGCGGTGAAGAAGA 60
QY 486 GCTTGACGGGCGCCCAACAGAGCTTCCTCAACAACCTTTTGTGCCCCGATTTTGGGA 545
DB 61 GCTTGACGGGCGCCCAACAGAGCTTCCTCAACAACCTTTTGTGCCCCGATTTTGGGA 120
QY 546 GTGATGATGAATCATGATCATGATGAATATCTTCTGCGGGGTCTTA 591
DB 121 GTGATGATGAATCATGATCATGATGAATATCTTCTGCGGGGTCTTA 166

Search completed: February 6, 2006, 10:41:01
Job time : 480 secs

This Page Blank (uspto)


```
OY 1 ATGAGGAGAGGCTTTGAAAGAGAAACACTCTGATTTTGGAAAGAGAGAGAGAGAGGTTGGAG 60
DB 1 ATGGTGGAGAGGCTTTGAAAGAGAAACACTCTGATTTTGGAAAGAGAGAGAGAGAGGTTGGAG 60
OY 61 TGGAGTCCAGAGAGGCTTTCAAGAGCTACAGAGAAAGCCCTGCTCCGATCTCCCTAGACAAA 120
DB 61 TGGAGTCCAGAGAGGCTTTCAAGAGCTACAGAGAAAGCCCTGCTCCGATCTCCCTAGACAAA 120
OY 121 GTCCAGCGAGGCTTGGAGCCGCCAGAGACCCAGCCTCCGAGGAGATGCTCTCATTCATAAC 180
DB 121 GTCCAGCGAGGCTTGGAGCCGCCAGAGACCCAGCCTCCGAGGAGATGCTCTCATTCATAAC 180
OY 181 ACCCTTCACAGAGCTGAGAGCTGACCTTGGCTGAGCTCCGCTCCGCTCCGCTCCGCTCCGAG 240
DB 181 ACCCTTCACAGAGCTGAGAGCTGACCTTGGCTGAGCTCCGCTCCGCTCCGCTCCGCTCCGAG 240
OY 241 CCCCTCTTCTGGAGGAGAGAGATTTCTCCCTGTCAGAGCAGCATTTGGCTCTATTCCTCAAG 300
DB 241 CCCCTCTTCTGGAGGAGAGAGATTTCTCCCTGTCAGAGCAGCATTTGGCTCTATTCCTCAAG 300
OY 301 GAGCTGACACCTTCATGATGAGAGCTGAGAGCCCTCCAGAAATCCAGTGACTCCCTTGGC 360
DB 301 GAGCTGACACCTTCATGATGAGAGCTGAGAGCCCTCCAGAAATCCAGTGACTCCCTTGGC 360
OY 361 CTCAGAAATGAAGTGCACCCCAAGCTGATCCAGTCTTTTGAAGAGCTGTAGCTCCG 420
DB 361 CTCAGAAATGAAGTGCACCCCAAGCTGATCCAGTCTTTTGAAGAGCTGTAGCTCCG 420
OY 421 TACTTGGGGGAGACTGAGCTGATGATGATCTTTCTTGGACATGACATCTGCGGTAGAA 480
DB 421 TACTTGGGGGAGACTGAGCTGATGATGATCTTTCTTGGACATGACATCTGCGGTAGAA 480
OY 481 AAGAGAGCTGACAGGAGCCCAAGAGCCTCTCAACAACCTTCTGAGCCCAAGTCT 540
DB 481 AAGAGAGCTGACAGGAGCCCAAGAGCCTCTCAACAACCTTCTGAGCCCAAGTCT 540
OY 541 TGGAGTGAATGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 591
DB 541 TGGAGTGAATGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 591
```

```
RESULT 2
CN304585 729 bp mRNA linear EST 16-MAY-2004
LOCUS DEFINITION 17000600512783 GRN_PRENEM Homo sapiens cDNA 5', mRNA sequence.
ACCESSION CN304585
VERSION CN304585.1 GI:47320999
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
1 (bases 1 to 729)
```

Bradenberger, R., Wei, H., Zhang, S., Lei, S., Muraue, J., Fisk, G.J.,
Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandalam, R.,
Lebowksi, J. and Stanton, L.W.

Transcriptome characterization elucidates signaling networks that
control human ES cell growth and differentiation

Nat. Biotechnol. 22 (6), 707-716 (2004)
15146197

Contact: Bradenberger R
Regenerative Medicine
Geron Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbradenberger@geron.com

Insert Length: 729 Std Error: 0.00.
Location/Qualifiers

1..729
/organism="Homo sapiens"
/mol_type="mRNA"

/db_xref="taxon:9606"
/tissue_type="embryonic stem cell, retinoic acid and
mitogen-treated hES cell line H7"
/clone_id="GRN_PRENEM"
/note="oligo dT primed, full-length enriched cDNA library
from hES cell line H7 (p29) maintained in feeder-free
conditions. Embryoid bodies were generated in the presence
of all-trans retinoic acid and mitogens."

Query Match 99.7%; Score 589.4; DB 7; Length 729;
Best Local Similarity 99.8%; Pred. No. 2,1e-135;
Matches 590; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
OY 1 ATGAGGAGAGGCTTTGAAAGAGAAACACTCTGATTTTGGAAAGAGAGAGAGAGGTTGGAG 60
DB 98 ATGGTGGAGAGGCTTTGAAAGAGAAACACTCTGATTTTGGAAAGAGAGAGAGAGGTTGGAG 157
OY 61 TGGAGTCCAGAGAGGCTTTCAAGAGCTACAGAGAAAGCCCTGCTCCGATCTCCCTAGACAAA 120
DB 158 TGGAGTCCAGAGAGGCTTTCAAGAGCTACAGAGAAAGCCCTGCTCCGATCTCCCTAGACAAA 217
OY 121 GTCCAGCGAGGCTTGGAGCCGCCAGAGACCCAGCCTCCGAGGAGATGCTCTCATTCATAAC 180
DB 218 GTCCAGCGAGGCTTGGAGCCGCCAGAGACCCAGCCTCCGAGGAGATGCTCTCATTCATAAC 277
OY 181 ACCCTTCACAGAGCTGAGAGCTGACCTTGGCTGAGCTCCGCTCCGCTCCGCTCCGCTCCGAG 240
DB 278 ACCCTTCACAGAGCTGAGAGCTGACCTTGGCTGAGCTCCGCTCCGCTCCGCTCCGCTCCGAG 337
OY 241 CCCCTCTTCTGGAGGAGAGAGATTTCTCCCTGTCAGAGCAGCATTTGGCTCTATTCCTCAAG 300
DB 338 CCCCTCTTCTGGAGGAGAGAGATTTCTCCCTGTCAGAGCAGCATTTGGCTCTATTCCTCAAG 397
OY 301 GAGCTGACACCTTCATGATGAGAGCTGAGAGCCCTCCAGAAATCCAGTGACTCCCTTGGC 360
DB 398 GAGCTGACACCTTCATGATGAGAGCTGAGAGCCCTCCAGAAATCCAGTGACTCCCTTGGC 457
OY 361 CTCAGAAATGAAGTGCACCCCAAGCTGATCCAGTCTTTTGAAGAGCTGTAGCTCCG 420
DB 458 CTCAGAAATGAAGTGCACCCCAAGCTGATCCAGTCTTTTGAAGAGCTGTAGCTCCG 517
OY 421 TACTTGGGGGAGACTGAGCTGATGATGATCTTTCTTGGACATGACATCTGCGGTAGAA 480
DB 518 TACTTGGGGGAGACTGAGCTGATGATGATCTTTCTTGGACATGACATCTGCGGTAGAA 577
OY 481 AAGAGAGCTGACAGGAGCCCAAGAGCCTCTCAACAACCTTCTGAGCCCAAGTCT 540
DB 578 AAGAGAGCTGACAGGAGCCCAAGAGCCTCTCAACAACCTTCTGAGCCCAAGTCT 637
OY 541 TGGAGTGAATGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 591
DB 638 TGGAGTGAATGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 688
```

RESULT 3
CN304584 742 bp mRNA linear EST 16-MAY-2004
LOCUS DEFINITION 17000597971761 GRN_PRENEM Homo sapiens cDNA 5', mRNA sequence.
ACCESSION CN304584
VERSION CN304584.1 GI:47320998
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
1 (bases 1 to 742)

Bradenberger, R., Wei, H., Zhang, S., Lei, S., Muraue, J., Fisk, G.J.,
Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandalam, R.,
Lebowksi, J. and Stanton, L.W.

Transcriptome characterization elucidates signaling networks that
control human ES cell growth and differentiation

JOURNAL Nat. Biotechnol. 22 (6), 707-716 (2004)
PUBMED 15146197
COMMENT Contact: Brandenberger R

ORIGIN

Query Match	99.7%	Score 589.4;	DB 7;	Length 742;
Best Local Similarity	99.8%	Pred. No. 2.1e-135;		
Matches 590; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0

RESULT 4	1038 bp	mRNA	linear	EST 20-FEB-2007
BM560308				
LOCUS				
DEFINITION				
	BM560308			
	AGNCENTR_6563953	NIH_MGC_119	Homo sapiens	cdna IMAGE:574771
	5', mRNA	sequence.		

ACCESSION	BM560308
VERSION	BM560308.1
KEYWORDS	GI:18804629
SOURCE	EST.
ORGANISM	Homo sapiens (human)
	Homo sapiens

REFERENCE	1 (bases 1 to 1038)
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC).
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D. rstraus@nsl.jhu.edu

FEATURES

SOURCE

000000

Query Match	99.7%	Score 589.4;	DB 3;	Length 1036;
Best Local Similarity	99.8%;	Pred. No. 2.2e-135;		
Matches 590;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;

DB	SEQUENCE	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	REMARK	COMMENT	FEATURES	ORIGIN
Db	CTCCGAGATGAAAGTGCACCCAGGCTATCCAGTCTTTGAGAGCTGTAGCTCCGG	CR605181	full-length cDNA clone CS0DJ012YJ05 of T cells (Jurkat cell line)	CR605181	GI:50485988	HTC; CNSLT cDNA.	HTC	Homo sapiens (human)	Genoscope.	Direct Submission	Submitted (20-JUN-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr)				1.1298	1.1298
Qy	TACTTGGGAGACTCTGGCCTGTGATGACTTTCTTCTGGAACATTGACACATCTGGGTAGA	CR605181	full-length cDNA clone CS0DJ012YJ05 of T cells (Jurkat cell line)	CR605181	GI:50485988	HTC; CNSLT cDNA.	HTC	Homo sapiens (human)	Genoscope.	Direct Submission	Submitted (20-JUN-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr)				1.1298	1.1298
Db	TACTTGGGAGACTCTGGCCTGTGATGACTTTCTTCTGGAACATTGACACATCTGGGTAGA	CR605181	full-length cDNA clone CS0DJ012YJ05 of T cells (Jurkat cell line)	CR605181	GI:50485988	HTC; CNSLT cDNA.	HTC	Homo sapiens (human)	Genoscope.	Direct Submission	Submitted (20-JUN-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr)				1.1298	1.1298
Qy	AAAGAGCCTGCACCGGGCCCAACCAAGCCTCTCAACAACCTCTTGTGGCCCAAGTCT	CR605181	full-length cDNA clone CS0DJ012YJ05 of T cells (Jurkat cell line)	CR605181	GI:50485988	HTC; CNSLT cDNA.	HTC	Homo sapiens (human)	Genoscope.	Direct Submission	Submitted (20-JUN-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr)				1.1298	1.1298
Db	AAAGAGCCTGCACCGGGCCCAACCAAGCCTCTCTCAACAACCTCTTGTGGCCCAAGTCT	CR605181	full-length cDNA clone CS0DJ012YJ05 of T cells (Jurkat cell line)	CR605181	GI:50485988	HTC; CNSLT cDNA.	HTC	Homo sapiens (human)	Genoscope.	Direct Submission	Submitted (20-JUN-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr)				1.1298	1.1298
Qy	TGGAGTGGAAATGAATGATGCATCATCATGAAATCATTTCTGGGCTCTAA	CR605181	full-length cDNA clone CS0DJ012YJ05 of T cells (Jurkat cell line)	CR605181	GI:50485988	HTC; CNSLT cDNA.	HTC	Homo sapiens (human)	Genoscope.	Direct Submission	Submitted (20-JUN-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr)				1.1298	1.1298
Db	TGGAGTGGAAATGAATGATGCATCATCATGAAATCATTTCTGGGCTCTAA	CR605181	full-length cDNA clone CS0DJ012YJ05 of T cells (Jurkat cell line)	CR605181	GI:50485988	HTC; CNSLT cDNA.	HTC	Homo sapiens (human)	Genoscope.	Direct Submission	Submitted (20-JUN-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr)				1.1298	1.1298
Qy	ATGAGAGGAGGCTTGAAGAAGAAACACTGTGATTTGAAAGAGAGAGAGAGTGGAG	CR605181	full-length cDNA clone CS0DJ012YJ05 of T cells (Jurkat cell line)	CR605181	GI:50485988	HTC; CNSLT cDNA.	HTC	Homo sapiens (human)	Genoscope.	Direct Submission	Submitted (20-JUN-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr)				1.1298	1.1298
Db	ATGAGAGGAGGCTTGAAGAAGAAACACTGTGATTTGAAAGAGAGAGAGAGTGGAG	CR605181	full-length cDNA clone CS0DJ012YJ05 of T cells (Jurkat cell line)	CR605181	GI:50485988	HTC; CNSLT cDNA.	HTC	Homo sapiens (human)	Genoscope.	Direct Submission	Submitted (20-JUN-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr)				1.1298	1.1298
Qy	TGGAGTCCAGCAGGCTTCAAGACTTACAGCAAGCCTTCTCGCATCTCCTTAACAA	CR605181	full-length cDNA clone CS0DJ012YJ05 of T cells (Jurkat cell line)	CR605181	GI:50485988	HTC; CNSLT cDNA.	HTC	Homo sapiens (human)	Genoscope.	Direct Submission	Submitted (20-JUN-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr)				1.1298	1.1298
Db	TGGAGTCCAGCAGGCTTCAAGACTTACAGCAAGCCTTCTCGCATCTCCTTAACAA	CR605181	full-length cDNA clone CS0DJ012YJ05 of T cells (Jurkat cell line)	CR605181	GI:50485988	HTC; CNSLT cDNA.	HTC	Homo sapiens (human)	Genoscope.	Direct Submission	Submitted (20-JUN-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr)				1.1298	1.1298
Qy	GTTCAGGCGAGCTGGGACCCCGAGAGACCCAGCTCCGAGGAGCATCTCATATCAAC	CR605181	full-length cDNA clone CS0DJ012YJ05 of T cells (Jurkat cell line)	CR605181	GI:50485988	HTC; CNSLT cDNA.	HTC	Homo sapiens (human)	Genoscope.	Direct Submission	Submitted (20-JUN-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr)				1.1298	1.1298
Db	GTTCAGGCGAGCTGGGACCCCGAGAGACCCAGCTCCGAGGAGCATCTCATATCAAC	CR605181	full-length cDNA clone CS0DJ012YJ05 of T cells (Jurkat cell line)	CR605181	GI:50485988	HTC; CNSLT cDNA.	HTC	Homo sapiens (human)	Genoscope.	Direct Submission	Submitted (20-JUN-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr)				1.1298	1.1298
Qy	GTTCAGGCGAGCTGGGACCCCGAGAGACCCAGCTCCGAGGAGCATCTCATATCAAC	CR605181	full-length cDNA clone CS0DJ012YJ05 of T cells (Jurkat cell line)	CR605181	GI:50485988	HTC; CNSLT cDNA.	HTC	Homo sapiens (human)	Genoscope.	Direct Submission	Submitted (20-JUN-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr)				1.1298	1.1298
Db	GTTCAGGCGAGCTGGGACCCCGAGAGACCCAGCTCCGAGGAGCATCTCATATCAAC	CR605181	full-length cDNA clone CS0DJ012YJ05 of T cells (Jurkat cell line)	CR605181	GI:50485988	HTC; CNSLT cDNA.	HTC	Homo sapiens (human)	Genoscope.	Direct Submission	Submitted (20-JUN-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr)				1.1298	1.1298

Qy	181	ACCTCCAAAGCTGACAGGCTGACCTTGACTCGCTCCCGCCCTGACCTCCCGAG	240
Db	285	ACCTCCAAAGCTGACAGGCTGACCTTGACTCGCTCCCGCCCTGACCTCCCGAG	344
Qy	241	CCCTCTTCTCGGGCGAGAGGATTTCTCCCTGTCAAGCAACATTGGCTTATCTGAG	300
Db	345	CCCTCTTCTCGGGCGAGAGGATTTCTCCCTGTCAAGCAACATTGGCTTATCTGAG	404
Qy	301	GAGCTGGACACTTCATGATGATGGATCTGAGCCCTCCAGATCCAGATCTCCCTGGC	360
Db	405	GAGCTGGACACTTCATGATGATGGATCTGAGCCCTCCAGATCCAGATCTCCCTGGC	464
Qy	361	CTCCAGATGAAGTGGCCACCCCGACCTGATCCAGTCTTCTTAAGAGCTTGAAGCTCCCG	420
Db	465	CTCCAGATGAAGTGGCCACCCCGACCTGATCCAGTCTTCTTAAGAGCTTGAAGCTCCCG	524
Qy	421	TACTTGGGGGACTCTGGCTTGAGTGAATCTTTCTTGACATTGACACATCTGCGGTGAA	480
Db	525	TACTTGGGGGACTCTGGCTTGAGTGAATCTTTCTTGACATTGACACATCTGCGGTGAA	584
Qy	481	AAGAGCTTGAAGGGGCCCCACAGAGCTTCTTCAAACTTCTGTGCCCCAGATTCT	540
Db	585	AAGAGCTTGAAGGGGCCCCACAGAGCTTCTTCAAACTTCTGTGCCCCAGATTCT	644
Qy	541	TGGAGTGAATGAATGATGATCATCATGTAATCATTTCTGGGGTCTTA	591
Db	645	TGGAGTGAATGAATGATGATCATCATGTAATCATTTCTGGGGTCTTA	695
RESULT 6			
LOCUS	CR601245	1318 bp	mRNA linear HTC 21-JUL-2004
DEFINITION	full-length cDNA clone CSOD1057YM14 of Placenta Cot 25-normalized		
ACCESSION	CR601245		
VERSION	CR601245.1	GI:50482052	
KEYWORDS	HTC; CNSLT cDNA.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Bukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 1 to 1318)		
TITLE	Li, W.B., Gruber, C., Jesse, J. and Polayes, D.		
JOURNAL	Full-length cDNA libraries and normalization		
REMARK	Unpublished		
REFERENCE	Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600		
AUTHORS	2 (bases 1 to 1318)		
TITLE	Faraday Avenue		
JOURNAL	Genoscope.		
COMMENT	Direct Submission		
FEATURES	Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 Evry cedex - FRANCES (E-mail : sequef@genoscope.cns.fr		
source	- Web : www.genoscope.cns.fr)		
	1st strand cDNA was primed with a NotI-oligo(dT) primer. Five primers		
	end enriched, double-strand cDNA was digested with Not I and cloned		
	into the Not I and Bclor V sites of the pCMVSPORT 6 vector. Library		
	was normalized. Library was constructed by Life Technologies, a		
	division of Invitrogen.		
	Location/Qualifiers		
	1..1318		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="CSOD1057YM14"		
	/tissue_type="Placenta Cot 25-normalized"		
	/plasmid="pCMVSPORT_6"		
ORIGIN			
Query Match	99.7%	Score 589.4	DB 4; Length 1318;
Best Local Similarity	99.8%	Pred. No. 2.3e-135;	

Matches 590; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAGGAGGCTTGAAGAGAAACCTGTGATTTGAAAGAGAGAGAGAGTGGAG 60
DB 112 ATGTGTGGAGGCTTGAAGAGAAACCTGTGATTTGAAAGAGAGAGAGTGGAG 171
QY 61 TGGAGTCCAGAGGCTTCAAGACTACAGCAAGCCCTGCTCCGATCTCCCTAGACAA 120
DB 172 TGGAGTCCAGAGGCTTCAAGACTACAGCAAGCCCTGCTCCGATCTCCCTAGACAA 231
QY 121 GTCCAGCCGAGCTGGGCCCCGAGACCCACCTCTCCGAGGAGATGTCTCATTCATAC 180
DB 232 GTCCAGCCGAGCTGGGCCCCGAGACCCACCTCTCCGAGGAGATGTCTCATTCATAC 291
QY 181 ACCCTCAACAGCTGACAGCTGACCTGAGCTGGCTCCGCCCCCTGCCCCCGAG 240
DB 292 ACCCTCAACAGCTGACAGCTGACCTGAGCTGGCTCCGCCCCCTGCCCCCGAG 351
QY 241 CCCCTCTTCTGGGCGAGAGGATTTCTCCCTGTGACGACCATTTGGCTTATCTCAAG 300
DB 352 CCCCTCTTCTGGGCGAGAGGATTTCTCCCTGTGACGACCATTTGGCTTATCTCAAG 411
QY 301 GAGCTGACACCTCCATGATGAGGACTGAGCCCTCCAGATCCAGTACTCCCTGGC 360
DB 412 GAGCTGACACCTCCATGATGAGGACTGAGCCCTCCAGATCCAGTACTCCCTGGC 471
QY 361 CTCGAATGAAGTGCACCCGAGCTGATCCAGTCTTTAGAAAGCTGTAGCTCCGG 420
DB 472 CTCGAATGAAGTGCACCCGAGCTGATCCAGTCTTTAGAAAGCTGTAGCTCCGG 531
QY 421 TACTTGGGGAAGCTGAGCTGTGATGATCTTCTTGTGACATTTGACATCTGCGTAGA 480
DB 532 TACTTGGGGAAGCTGAGCTGTGATGATCTTCTTGTGACATTTGACATCTGCGTAGA 591
QY 481 AAGAGCCTGACGCGGCCCCACCAAGGCTCTCTCAACACTTCTTGTGCGCCAGATTCT 540
DB 592 AAGAGCCTGACGCGGCCCCACCAAGGCTCTCTCAACACTTCTTGTGCGCCAGATTCT 651
QY 541 TGGGAGTGAATGAAGTGAATGATCATCATGAAATCATTTGCGGCTCTTA 591
DB 652 TGGGAGTGAATGAAGTGAATGATCATCATGAAATCATTTGCGGCTCTTA 702

RESULT 7
BU157342 1056 bp mRNA linear EST 04-SEP-2002
LOCUS AGENCOURT_6925371 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:5952523
DEFINITION BU157342 5', mRNA sequence.
ACCESSION BU157342.1 GI:22670874
VERSION EST.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homniidae; Homo.
1 (bases 1 to 1056)
NIH-MGC <http://mgc.nhi.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Place: LDCM2140 row: k column: 20
High quality sequence stop: 587.
Location/Qualifiers
1. 1056

FEATURES
source

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5952523"
/tissue_type="ductal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/name_1="NIH_MGC_110"
/note="Organ: pancreas; Vector: pOTB7; Site: 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGGACAG(C). Library constructed by
Ling Hong in the Laboratory of Gerald M. Rubin (University
of California, Berkeley) using Zap-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN

Query Match 98.6%; Score 582.6; DB 5; Length 1056;
Best Local Similarity 99.0%; Pred. No. 1.1e-133;
Matches 585; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 ATGAGGAGGCTTGAAGAGAAACCTGTGATTTGAAAGAGAGAGAGTGGAG 60
DB 108 ATGTGTGGAGGCTTGAAGAGAAACCTGTGATTTGAAAGAGAGAGTGGAG 167
QY 61 TGGAGTCCAGAGGCTTCAAGACTACAGCAAGCCCTGCTCCGATCTCCCTAGACAA 120
DB 168 TGGAGTCCAGAGGCTTCAAGACTACAGCAAGCCCTGCTCCGATCTCCCTAGACAA 227
QY 121 GTCCAGCCGAGCTGGGCCCCGAGACCCACCTCTCCGAGGAGATGTCTCATTCATAC 180
DB 228 GTCCAGCCGAGCTGGGCCCCGAGACCCACCTCTCCGAGGAGATGTCTCATTCATAC 287
QY 181 ACCCTCAACAGCTGACAGCTGACCTGAGCTGGCTCCGCCCCCTGCCCCCGAG 240
DB 288 ACCCTCAACAGCTGACAGCTGACCTGAGCTGGCTCCGCCCCCTGCCCCCGAG 347
QY 241 CCCCTCTTCTGGGCGAGAGGATTTCTCCCTGTGACGACCATTTGGCTTATCTCAAG 300
DB 348 CCCCTCTTCTGGGCGAGAGGATTTCTCCCTGTGACGACCATTTGGCTTATCTCAAG 407
QY 301 GAGCTGACACCTCCATGATGAGGACTGAGCCCTCCAGATCCAGTACTCCCTGGC 360
DB 408 GAGCTGACACCTCCATGATGAGGACTGAGCCCTCCAGATCCAGTACTCCCTGGC 467
QY 361 CTCGAATGAAGTGCACCCGAGCTGATCCAGTCTTTAGAAAGCTGTAGCTCCGG 420
DB 468 CTCGAATGAAGTGCACCCGAGCTGATCCAGTCTTTAGAAAGCTGTAGCTCCGG 527
QY 421 TACTTGGGGAAGCTGAGCTGTGATGATCTTCTTGTGACATTTGACATCTGCGTAGA 480
DB 528 TACTTGGGGAAGCTGAGCTGTGATGATCTTCTTGTGACATTTGACATCTGCGTAGA 587
QY 481 AAGAGCCTGACGCGGCCCCACCAAGGCTCTCTCAACACTTCTTGTGCGCCAGATTCT 540
DB 588 AAGAGCCTGACGCGGCCCCACCAAGGCTCTCTCAACACTTCTTGTGCGCCAGATTCT 647
QY 541 TGGGAGTGAATGAAGTGAATGATCATCATGAAATCATTTGCGGCTCTTA 591
DB 648 TGGGAGTGAATGAAGTGAATGATCATCATGAAATCATTTGCGGCTCTTA 698

RESULT 8
CN304590 646 bp mRNA linear EST 16-MAY-2004
LOCUS CN304590 17000600179006 GRN_PRENEMU Homo sapiens cDNA 5', mRNA sequence.
DEFINITION CN304590
ACCESSION CN304590.1 GI:47321004
VERSION EST.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

REFERENCE 1 (bases 1 to 646)
 AUTHORS Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G., Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M. S., Mandalam, R., Lebowicki, J. and Stanton, L. W.
 TITLE Transcriptome characterization elucidates signaling networks that control human ES cell growth and differentiation
 JOURNAL Nat. Biotechnol. 22 (6), 707-716 (2004)
 PUBMED 15146197
 COMMENT Contact: Brandenberger R
 Regenerative Medicine
 Genon Corporation
 230 Constitution Drive, Menlo Park, CA 94025, USA
 Tel: 650 473 8658
 Fax: 650 473 7760
 Email: rbrandenberger@genon.com
 Insert Length: 646 Std Error: 0.00.
 Location/Qualifiers
 1..646
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /tissue_type="embryonic stem cell, retinoic acid and mitogen-treated hES cell line H7"
 /note="Oligo dt primed, full-length enriched cDNA library from hES cell line H7 (p29) maintained in feeder-free conditions. Embryoid bodies were generated in the presence of all-trans retinoic acid and mitogens."

ORIGIN

Query Match 98.2%; Score 580.4; DB 7; Length 646;
 Best Local Similarity 99.8%; Pred. No. 3.5e-133;
 Matches 581; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 ATGAGGAGAGGCTTGAAGAGAAACCTGTGATTGAGAGAGAGAGAGAGTGGAG 60
DB 65 ATGAGGAGAGGCTTGAAGAGAAACCTGTGATTGAGAGAGAGAGAGAGTGGAG 124
QY 61 TGGAGTCCAGCAGGCTTCAAGACTACAGCAAGCCTGCTCCGATCTCCCTAGACAA 120
DB 125 TGGAGTCCAGCAGGCTTCAAGACTACAGCAAGCCTGCTCCGATCTCCCTAGACAA 184
QY 121 GTCCAGCAGCCTGGGCCCCGAGACCCAGCCTCCGAGAGATGCTCATTCATAAC 180
DB 185 GTCCAGCAGCCTGGGCCCCGAGACCCAGCCTCCGAGAGATGCTCATTCATAAC 244
QY 181 ACCCTCCACAGCTGAGAGCTGACATTCGCTGGCTCCGCCCCCTGCCCCCGAG 240
DB 245 ACCCTCCACAGCTGAGAGCTGACATTCGCTGGCTCCGCCCCCTGCCCCCGAG 304
QY 241 CCCCTCTTCTGGGCGAGAGATTCTCCCTGTGAGCACCATTGGCTCTATCTCAG 300
DB 305 CCCCTCTTCTGGGCGAGAGATTCTCCCTGTGAGCACCATTGGCTCTATCTCAG 364
QY 301 GAGCTGACACCTCCATGATGAGATGAGCCCCCTCAGAAATCCAGTGACTCCCTTGC 360
DB 365 GAGCTGACACCTCCATGATGAGATGAGCCCCCTCAGAAATCCAGTGACTCCCTTGC 424
QY 361 CTCGAGATGAAAGTGCACCCAGCTGATTCAGATTCTTTAGAAAGCTCTGAGCTCCG 420
DB 425 CTCGAGATGAAAGTGCACCCAGCTGATTCAGATTCTTTAGAAAGCTCTGAGCTCCG 484
QY 421 TACTTGGGGGAGCTTGGGCTGAGATGCTTTCTTGGAGATTGACATCTGCGGTGAA 480
DB 485 TACTTGGGGGAGCTTGGGCTGAGATGCTTTCTTGGAGATTGACATCTGCGGTGAA 544
QY 481 AAGAGCCTGACAGGCCCCCAGACAGCCTCTCACAACCTCTTGTGCCCCCAGATTCT 540
DB 545 AAGAGCCTGACAGGCCCCCAGACAGCCTCTCACAACCTCTTGTGCCCCCAGATTCT 604
QY 541 TGGAGTGAATGAATTGATCATCATGAAATCATTTCTG 582

```

DB 605 TGGAGTGAATGAATGACTGATCATCATGAAATCATTTCTG 646

RESULT 9
 EX398356
 LOCUS
 DEFINITION BX398356 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA clone CS0D1057YM14 5-PRIME, mRNA sequence.
 ACCESSION BX398356
 VERSION BX398356
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1017)
 Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 On May 13, 2003 this sequence version replaced gi:30613656.
 CONTACT Genoscope
 Genoscope - Centre National de Sequencage
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo(dt) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and BcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 9131.r

For more information about this cluster, see
 http://www.genoscope.cns.fr/cdna78/CS0D1057BG07QPlc=9131.r.
 Location/Qualifiers
 1..1017
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0D1057YM14"
 /tissue_type="PLACENTA COT 25-NORMALIZED"
 /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dt) primer. Five prime end enriched, double-strand cDNA was digested with Not I and BcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 97.4%; Score 575.8; DB 5; Length 1017;
 Best Local Similarity 99.5%; Pred. No. 5.2e-132;
 Matches 588; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

```

QY 1 ATGAGGAGAGGCTTGAAGAGAAACCTGTGATTGAGAGAGAGAGAGTGGAG 60
DB 112 ATGAGGAGAGGCTTGAAGAGAAACCTGTGATTGAGAGAGAGAGAGTGGAG 171
QY 61 TGGAGTCCAGCAGGCTTCAAGACTACAGCAAGCCTGCTCCGATCTCCCTAGACAA 120
DB 172 TGGAGTCCAGCAGGCTTCAAGACTACAGCAAGCCTGCTCCGATCTCCCTAGACAA 231
QY 121 GTCCAGCAGCCTGGGCCCCGAGACCCAGCCTCCGAGAGATGCTCATTCATAAC 180
DB 232 GTCCAGCAGCCTGGGCCCCGAGACCCAGCCTCCGAGAGATGCTCATTCATAAC 290
QY 222 GTCCAGCAGCCTGGGCCCCGAGACCCAGCCTCCGAGAGATGCTCATTCATAAC 240
DB 291 ACCCTCCACAGCTGAGAGCTGACATTCGCTGGCTCCGCCCCCTGCCCCCGAG 350
QY 241 CCCCTCTTCTGGGCGAGAGATTCTCCCTGTGAGCACCATTGGCTCTATCTCAG 300
DB 351 CCCCTCTTCTGGGCGAGAGATTCTCCCTGTGAGCACCATTGGCTCTATCTCAG 410
QY 301 GAGCTGACACCTCCATGATGAGATGAGCCCCCTCAGAAATCCAGTGACTCCCTTGC 360

```


Db 411 GAGCTGACACCTCCATGATGAGGAGTGGACGCCCCCTCAGATTCAGTGAATCTCCCTTGGC 470
Qy 361 CTCGAGAAATGAAGTGCACCCCGAGCTGATCCAGTCTTTAGAAAGCTTGAGCTCCCG 420
Db 471 CTCGAGAAATGAAGTGCACCCCGAGCTGATCCAGTCTTTAGAAAGCTTGAGCTCCCG 530
Qy 421 TACTTGGGGAGACTGGGCTGATGATCTTTCTTGGACATTTGACATCTGCGGTAGA 480
Db 531 TACTTGGGGAGACTGGGCTGATGATCTTTCTTGGACATTTGACATCTGCGGTAGA 590
Qy 481 AAGAGGCTGACGCGGCCCGACCAAGAGCTCTCTCACAACCTCTTGTGCCCCAGGTTCT 540
Db 591 AAGAGGCTGACGCGGCCCGACCAAGAGCTCTCTCACAACCTCTTGTGCCCCAGGTTCT 650
Qy 541 TGGGAGTGAATGAATGATGATCATCATGAAATCATCTTGGGGTCTTAA 591
Db 651 TGGGAGTGAATGAATGATGATCATCATGAAATCATCTTGGGGTCTTAA 701

RESULT 10

LOCUS CN304594 712 bp mRNA linear EST 16-MAY-2004
DEFINITION 17000532611443 GRN_BB Homo sapiens cDNA 5', mRNA sequence.

ACCESSION CN304594

VERSION CN304594.1 GI:47321008

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
1 (bases 1 to 712)

REFERENCE

AUTHORS

Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J.,
Li, Y., Xu, C., Pang, R., Guegler, K., Rao, M.S., Mandalam, R.,
Lebkowski, J. and Stanton, L.W.
Transcriptome characterization elucidates signaling networks that
control human ES cell growth and differentiation

JOURNAL

Nat. Biotechnol. 22 (6), 707-716 (2004)

PUBMED

COMMENT

Contact: Brandenberger R
Regenerative Medicine
Geron Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@geron.com

Insert Length: 712 Std Error: 0.00.
Location/Qualifiers

FEATURES

source

1..712
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic stem cells, embryoid bodies
derived from H1, H7 and H9 cells"
/clone_lib="GRN BB"
/note="Toiigo dt primed, full-length enriched cDNA library
from embryoid body outgrowths derived from h9s cell lines
H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free
conditions."

ORIGIN

Query Match 97.3%; Score 575.2; DB 7; Length 712;
Best Local Similarity 99.3%; Pred. No. 6.9e-132;
Matches 588; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy 1 ATGAGGAGGAGCTTGAAGAGAAACCTGTGATTGGAAGAGAGAGAGAGTGGAG 60
Db 115 ATGAGGAGGAGCTTGAAGAGAAACCTGTGATTGGAAGAGAGAGAGAGTGGAG 174
Qy 61 TGGAGTCCGACGAGGCTTCAAGACTACGACGAGAGCCCTGCTCCGATCTCCCTAGACAA 120
Db 175 TGGAGTCCGACGAGGCTTCAAGACTACGACGAGAGCCCTGCTCCGATCTCCCTAGACAA 234

Qy 121 GTCAGCCGAGCTTGGGCCCCCGAGACCCAGCTTCCGAGGAGATGTCTCATCCATTAAC 180
Db 225 GTCAGCCGAGCTTGGGCCCCCGAGACCCAGCTTCCGAGGAGATGTCTCATCCATTAAC 294
Qy 181 ACCCTCCAAACAGCTGACAGCTGACCTTGGGCTCCCGCCCTGCCCCCGAG 240
Db 235 ACCCTCCAAACAGCTGACAGCTGACCTTGGGCTCCCGCCCTGCCCCCGAG 354
Qy 241 CCCCTCTTCTGGGAGAGAGATTTCTCCCTGTAGACCAATTTGGCTCTAATCTCAG 300
Db 355 CCCCTCTTCTGGGAGAGAGATTTCTCCCTGTAGACCAATTTGGCTCTAATCTCAG 414
Qy 301 GAGCTGACACCTCCATGATGATGGAATGAGCCCTCAGAAATCCAGTCACTCCCTTGGC 360
Db 415 GAGCTGACACCTCCATGATGATGGAATGAGCCCTCAGAAATCCAGTCACTCCCTTGGC 474
Qy 361 CTCGAGAAATGAAGTGCACCCCGAGCTGATCCAGTCTTTTGAAGCTTGAGCTCCCG 420
Db 475 CTCGAGAAATGAAGTGCACCCCGAGCTGATCCAGTCTTTTGAAGCTTGAGCTCCCG 534
Qy 421 TACTTGGGGAGACTGGGCTGATGATCTTTCTTGGACATTTGACATCTGCGGTAGA 480
Db 535 TACTTGGGGAGACTGGGCTGATGATCTTTCTTGGACATTTGACATCTGCGGTAGA 594
Qy 481 AAGAGGCTGACGCGGCCCGACCAAGAGCTCTCTCACAACCTCTTGTGCCCCAGGTT- 539
Db 595 AAGAGGCTGACGCGGCCCGACCAAGAGCTCTCTCACAACCTCTTGTGCCCCAGGTTCC 654
Qy 540 TGGGAGTGAATGAATGATGATCATCATGAAATCATCTTGGGGTCTTAA 591
Db 655 TGGGAGTGAATGAATGATGATCATCATGAAATCATCTTGGGGTCTTAA 706

RESULT 11

B0508982

LOCUS

DEFINITION AGENCOURT 10095430 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:6503515
5', mRNA sequence.

ACCESSION B0508982

VERSION B0508982.1 GI:22815215

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
1 (bases 1 to 873)

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Robert Strauberg, Ph.D.
Email: cgapbe-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: L14M14061 row: e column: 20
High quality sequence stop: 605.
Location/Qualifiers

FEATURES

source

1..873
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_image="IMAGE:6503515"
/tissue_type="telomysarcoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_71"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2.1 kb."

ORIGIN

Query Match 96.1%; Score 567.8; DB 5; Length 873;
 Best Local Similarity 98.5%; Pred. No. 4,9e-130;
 Matches 583; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

1 ATGAGGAGAGGCTTGAAGAGAAACATCTGATTTTGGAAAGAGAGAGAGAGAGGCTGGAG 60
 179 ATGGTGGAGAGCTTGAAGAGAAACATCTGATTTTGGAAAGAGAGAGAGAGAGGCTGGAG 238
 61 TGGAGTCCAGAGAGGCTTCAAGAGTACAGAGAGAGGCTTCCGATCTCCCTAGACAAA 120
 239 TGGAGTCCAGAGAGGCTTCAAGAGTACAGAGAGAGGCTTCCGATCTCCCTAGACAAA 298
 121 GTCCAGCAGAGCTTGGGAG 180
 299 GTCCAGCAGAGCTTGGGAG 358
 181 ACCCTCCAAAGAGCTGAG 240
 359 ACCCTCCAAAGAGCTGAG 418
 241 CCCCTCTTCTGGGAG 300
 419 CCCCTCTTCTGGGAG 478
 301 GAGCTGACACCTCCATGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
 479 GAGCTGACACCTCCATGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 538
 361 CTCAGAGATGAAGTCCACCCAGCTGATCCAGTCTTGAAGAGCTTGAAGCTCCGAG 420
 539 CTCAGAGATGAAGTCCACCCAGCTGATCCAGTCTTGAAGAGCTTGAAGCTCCGAG 598
 421 TACTTGGGAGAGCTTGGGAG 480
 599 TACTTGGGAGAGCTTGGGAG 658
 481 AAGAGAGCTGACAG 539
 659 AAGAGAGCTGACAG 718
 540 TTGGAGATGAAGATGAAGATGAAGATGAAGATGAAGATGAAGATGAAGATGAAGATGA 591
 719 TTGGAGATGAAGATGAAGATGAAGATGAAGATGAAGATGAAGATGAAGATGAAGATGA 770

RESULT 12

COS79535

LOCUS

DEFINITION

COS79535 970 bp mRNA linear EST 20-JUL-2004
 ILUMIGEN MCQ 49981 Katze mmlt Macaca mulatta cDNA clone
 IBIM:19438 5' similar to Bases 114 to 970 highly similar to human
 SERRAD3 (hs.169138), mRNA sequence.

ACCESSION

COS79535.1 GI:50410495

VERSION

EST.

KEYWORDS

Macaca mulatta (rhesus monkey)
 Macaca mulatta
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Cercopithecoidea; Cercopithecinae; Macaca.

SOURCE

EST.

REFERENCE

AUTHORS

1 (bases 1 to 970)
 Magness, C.L., Fellin, P.C., Thomas, M.J., Korth, M.J., Agay, M.B.,
 Prohl, S.C., Fitzgibbon, M., Scherer, C.A., Miner, D.G., Katze, M.G. and
 Iadonato, S.P.

TITLE

Analysis of the Macaca mulatta transcriptome and the sequence
 divergence between Macaca and human

JOURNAL

Genome Biol. 6 (7), R60 (2005)

PUBMED

1598449

COMMENT

Contact: C. Magness
 Illumigen Biosciences Inc.
 2203 Airport Way S, Suite 450, Seattle, WA 98134, USA
 Tel: 2063780400

Fax: 2063780408

Email: cmagness@illumigen.com

Sequenced on 2004.07.15. 648 Q20 bases. Library Preparation: Prof.
 Michael Katze Lab at University of Washington DNA Sequencing:

Illumigen Biosciences Inc. For further information, see
<http://www.macaque.org>

PCR primers

FORWARD: CCTTCATTAAGGACAAA

BACKWARD: CACTATGGGCGAATGGTA

Insert Length: 970 SLD Error: 0.00

Plate: C1000412 row: B column: 10

Seq primer: CCTTCATTAAGGACAAA

POLYA-No.

FEATURES

source

1..970

/organism="Macaca mulatta"

/mol_type="mRNA"

/strain="indian"

/db_xref="taxon:9544"

/clone="IBIM:19438"

/sex="male"

/dev_stage="adult"

/lab_host="Electromax DH10B"

/clone_lib="Katze mmlt"

/note="Organ: Illum; Vector: pDONR 222; Site 1: Berg 1;
 Site 2: Berg 1; Created from Cloneminer cDNA Library
 Construction kit (catalog #18249-029)"

ORIGIN

Query Match 94.9%; Score 560.8; DB 7; Length 970;
 Best Local Similarity 97.8%; Pred. No. 2.8e-128;
 Matches 579; Conservative 0; Mismatches 12; Indels 1; Gaps 1;

1 ATGAGGAGAGGCTTGAAGAGAAACATCTGATTTTGAAGAGAGAGAGAGAGGTTGGAG 60
 121 ATGGTGGAGAGCTTGAAGAGAAACATCTGATTTTGAAGAGAGAGAGAGAGGTTGGAG 180
 61 TGGAGTCCAGAGAGGCTTCAAGAGTACAGAGAGAGGCTTCCGATCTCCCTAGACAAA 120
 181 TGGAGTCCAGAGAGGCTTCAAGAGTACAGAGAGAGGCTTCCGATCTCCCTAGACAAA 240
 121 GTCCAGGAGAGCTTGGGAG 180
 241 GTCCAGGAGAGCTTGGGAG 300
 181 ACCCTCCAAAGAGCTGAG 240
 301 ACCCTCCAAAGAGCTGAG 360
 241 CCCCTCTTCTGGGAG 300
 361 CCCCTCTTCTGGGAG 420
 301 GAGCTGACACCTCCATGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
 421 GAGCTGACACCTCCATGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
 361 CTCAGAGATGAAGTCCACCCAGCTGATCCAGTCTTGAAGAGCTTGAAGCTCCGAG 420
 481 CTCAGAGATGAAGTCCACCCAGCTGATCCAGTCTTGAAGAGCTTGAAGCTCCGAG 540
 421 TACTTGGGAGAGCTTGGGAG 480
 541 TACTTGGGAGAGCTTGGGAG 600
 481 AAGAGAGCTGACAG 540
 601 AAGAGAGCTTGGGAG 660
 541 TGGAGATGGG-AATGAATGATCATCATGAGAAATCTTGGGGTCTTAA 591
 661 TGGAGATGGAAATGAATGATCATCATGAGAAATCTTGGGGTCTTAA 712


```
Db 105 ATGTGGAGAGGCTTGAAGAGAAACCTCTGATTTTGAAGAGAGAGAGAGGTGGAG 164
Oy 61 TGGAGTCCAGAGGCTTCAAGACTACAGAGAGAGGCTGTCTCCGATCTCCCTAGACAAA 120
Db 165 TGGAGTCCAGAGGCTTCAAGACTACAGAGAGAGGCTGTCTCCGATCTCCCTAGACAAA 224
Oy 121 GTCCAGCGCAGCTTGGGCCCCCGAGACCCAGCCTCCGAGGAGATGCTTCATTCATAAC 180
Db 225 GTCCAGCGCAGCTTGGGCCCCCGAGACCCAGCCTCCGAGGAGATGCTTCATTCATAAC 283
Oy 181 ACCCTCCAAAGCTTGCAGCTTGTGCGCTGCGCTCCGCGCCCTGCGCCCGCCGAG 240
Db 284 ACCCTCCAAAGCTTGCAGCTTGTGCGCTGCGCTCCGCGCCCTGCGCCCGCCGAG 343
Oy 241 CCCCTCTTCTGGGAGAGAGATTTTCTCCCTGTGAGCAGCAGCATTTGGCTTATCTCAG 300
Db 344 CCCCTCTTCTGGGAGAGAGATTTTCTCCCTGTGAGCAGCAGCATTTGGCTTATCTCAG 403
Oy 301 GAGCTGACACCTCCATGATGAGAGCTGAGAGCCCTCCAGAAATCAAGTACTCCCTTGGC 360
Db 404 GAGCTGACACCTCCATGATGAGAGCTGAGAGCCCTCCAGAAATCAAGTACTCCCTTGGC 463
Oy 361 CTCCAGAAATGAAGTGCACCCAGCCTGATCCAGTCTTTTGAAGAGCTTGAAGCTCCGG 420
Db 464 CTCCAGAAATGAAGTGCACCCAGCCTGATCCAGTCTTTTGAAGAGCTTGAAGCTCCGG 523
Oy 421 TACTTGGGAGAGCTTGGGAGAGATTTTCTCCCTGTGAGCAGCAGCATTTGGCTTATCTCAG 480
Db 524 TACTTGGGAGAGCTTGGGAGAGATTTTCTCCCTGTGAGCAGCAGCATTTGGCTTATCTCAG 583
Oy 481 AAGAGCCTGACAGGAGCCCGACAGAGCCTCTCAAACTCTTGTGAGCCCGAGTTCT 540
Db 584 AAGAGCCTGACAGGAGCCCGACAGAGCCTCTCAAACTCTTGTGAGCCCGAGTTCT 643
Oy 541 TG-GAGATGGAATGAATGATGATCAATCATTTGAAGATCATTTCTGGG 585
Db 644 TGCGAGTGGAGATGAATGATGATCAATCATTTCTGGG 689

RESULT 15
BM546306 1045 bp mRNA linear Bst 20-FEB-2002
LOCUS AGENCOURT 6498625 NIH_MGC_125 Homo sapiens cDNA IMAGE:5722856
DEFINITION BM546306 5', mRNA sequence.
ACCESSION BM546306.1 GI:18779156
VERSION Bst.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 1045)
NIH-MGC http://mhc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Invitrogen
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/ILNL at:
http://image.llnl.gov
Plate: L1AM12709 row: j column: 09
High quality sequence start: 20
High quality sequence stop: 626.
Location/Qualifiers
1..1045
/organism="Homo sapiens"
/mol_type="mRNA"
```

```
/db_xref="taxon:9606"
/clone="IMAGE:5722856"
/lab_host="DH10B"
/clone_id="NIH_MGC_125"
/notes="Organ: ovary (pool of 3); Vector: pCMV-Sport6;
Site 1: EcoRV (destroyed); Site 2: NotI; RNA source pool
of three ovaries, from females ranging in age from 38 to
49 yo. Library is oligo-dT primed and directionally cloned
(EcoRV site is destroyed upon cloning). Average insert
size 2.1 kb, insert size range 1-3.5 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 036."

ORIGIN
Query Match 94.2%; Score 557; DB 3; Length 1045;
Best Local Similarity 98.6%; Pred. No. 2.5e-127;
Matches 582; Conservative 0; Mismatches 6; Indels 2; Gaps 2;

Oy 1 ATGAGAGAGGCTTGAAGAGAAACCTCTGATTTTGAAGAGAGAGAGAGGTGGAG 60
Db 208 ATGTGGAGAGGCTTGAAGAGAAACCTCTGATTTTGAAGAGAGAGAGAGGTGGAG 267
Oy 61 TGGAGTCCAGAGGCTTCAAGACTACAGAGAGAGGCTGTCTCCGATCTCCCTAGACAAA 120
Db 268 TGGAGTCCAGAGGCTTCAAGACTACAGAGAGAGGCTGTCTCCGATCTCCCTAGACAAA 327
Oy 121 GTCCAGCGCAGCTTGGGCCCCCGAGACCCAGCCTCCGAGGAGATGCTTCATTCATAAC 180
Db 328 GTCCAGCGCAGCTTGGGCCCCCGAGACCCAGCCTCCGAGGAGATGCTTCATTCATAAC 387
Oy 181 ACCCTCCAAAGCTTGCAGCTTGTGCGCTGCGCTCCGCGCCCTGCGCCCGCCGAG 240
Db 388 ACCCTCCAAAGCTTGCAGCTTGTGCGCTGCGCTCCGCGCCCTGCGCCCGCCGAG 447
Oy 241 CCCCTCTTCTGGGAGAGAGATTTTCTCCCTGTGAGCAGCAGCATTTGGCTTATCTCAG 300
Db 448 CCCCTCTTCTGGGAGAGAGATTTTCTCCCTGTGAGCAGCAGCATTTGGCTTATCTCAG 507
Oy 301 GAGCTGACACCTCCATGATGAGAGCTGAGAGCCCTCCAGAAATCAAGTACTCCCTTGGC 360
Db 508 GAGCTGACACCTCCATGATGAGAGCTGAGAGCCCTCCAGAAATCAAGTACTCCCTTGGC 567
Oy 361 CTCCAGAAATGAAGTGCACCCAGCCTGATCCAGTCTTTTGAAGAGCTTGAAGCTCCGG 420
Db 568 CTCCAGAAATGAAGTGCACCCAGCCTGATCCAGTCTTTTGAAGAGCTTGAAGCTCCGG 627
Oy 421 TACTTGGGAGAGCTTGGGAGAGATTTTCTCCCTGTGAGCAGCAGCATTTGGCTTATCTCAG 480
Db 628 TACTTGGGAGAGCTTGGGAGAGATTTTCTCCCTGTGAGCAGCAGCATTTGGCTTATCTCAG 687
Oy 481 AAGAGCCTGACAGGAGCCCGACAGAGCCTCTCAAACTCTTGTGAGCCCGAGTTCT 540
Db 688 AAGAGCCTGACAGGAGCCCGACAGAGCCTCTCAAACTCTTGTGAGCCCGAGTTCT 747
Oy 541 TGGAGATGG-AAATGAATGATGATCAATCATTTGAAGATCATTTCTGGGCTCC 588
Db 748 TGGAGATGGCAATGATGATGATCAATCATTTCTGGGCTCC 797

Search completed: February 6, 2006, 12:37:27
Job time : 3761 secs
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

CM nucleic - nucleic search, using SW model

Run on: February 6, 2006, 11:20:09 ; Search time 172 Seconds
(without alignments)
6107.785 Million cell updates/sec

Title: US-10-069-386A-1

Perfect score: 591
Sequence: 1 atggaggaggagcttgaagag.....aatcatctcgggtcctaa 591

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents, NA: *
1: /cgn2_6/ptodata/1/ina/1_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/H_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
7: /cgn2_6/ptodata/1/ina/PP_COMB.seq:*
8: /cgn2_6/ptodata/1/ina/RG_COMB.seq:*
9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	49.2	8.3	7218	2	US-08-232-463-14
2	47.4	8.0	2082	3	US-10-076-069-1
3	46.4	7.9	44952	3	US-09-949-016-12197
4	46.4	7.9	44960	3	US-09-949-016-15853
5	46.4	7.9	51552	3	US-09-733-294A-30
6	42.6	7.2	1603	3	US-10-000-489-71
7	41	6.9	1926	3	US-09-249-585A-2
8	41	6.9	1926	3	US-09-410-399-3
9	41	6.9	2580	3	US-09-050-863-2
10	41	6.9	2580	3	US-09-359-081-2
11	41	6.9	5452	2	US-09-130-114-1
12	41	6.9	8705	2	US-09-647-344A-14
13	41	6.9	9600	3	US-08-910-647-1
14	41	6.9	9600	3	US-09-620-925-1
15	41	6.9	10596	2	US-07-884-811-15
16	41	6.9	10596	2	US-07-885-971-15
17	41	6.9	10596	2	US-08-087-783A-15
18	41	6.9	10596	2	US-08-194-088B-15
19	41	6.9	10596	2	US-08-194-087-15
20	41	6.9	10596	6	PCT-US93-04648-15
21	41	6.9	16080	3	US-09-724-566A-48
22	41	6.9	16080	3	US-09-471-669A-48
23	40.8	6.9	50937	3	US-09-428-517-1
24	40.6	6.9	396	2	US-07-872-678A-9

C	25	40.4	6.8	2097	3	US-08-941-445A-10	Sequence 10, Appl
	26	40.2	6.8	2254	3	US-09-016-434-1126	Sequence 1126, Ap
	27	40.2	6.8	2573	2	US-08-884-681-2	Sequence 2, Appl
	28	40.2	6.8	2573	3	US-09-258-643-2	Sequence 2, Appl
	29	40.2	6.8	4403765	3	US-09-103-840A-2	Sequence 2, Appl
	30	40.2	6.8	441529	3	US-09-103-840A-1	Sequence 1, Appl
	31	40	6.8	2127	3	US-09-902-540-8358	Sequence 8358, Ap
	32	40	6.8	7628	3	US-09-902-540-875	Sequence 875, App
	33	39.6	6.7	1432	3	US-09-620-312D-420	Sequence 420, App
	34	39.6	6.7	2825	3	US-09-196-397-5	Sequence 5, Appl
	35	39.6	6.7	2825	2	US-09-952-677-5	Sequence 5, Appl
	36	39.6	6.7	4394	2	US-08-093-734-1	Sequence 1, Appl
	37	39.6	6.7	4394	2	US-08-444-623-1	Sequence 1, Appl
	38	39.6	6.7	4394	3	US-08-471-863-1	Sequence 1, Appl
	39	39.6	6.7	4394	3	US-09-342-563-1	Sequence 1, Appl
	40	39.6	6.7	4394	6	PCT-US94-08267-1	Sequence 1, Appl
	41	39.4	6.7	780	3	US-09-902-540-4017	Sequence 4017, Ap
	42	39.4	6.7	2481	3	US-09-894-998A-35	Sequence 35, Appl
	43	39.4	6.7	2481	3	US-10-237-551-35	Sequence 35, Appl
	44	39.4	6.7	3066	3	US-10-237-551-152	Sequence 152, App
	45	39.4	6.7	26533	3	US-09-902-540-1199	Sequence 1199, Ap

ALIGNMENTS

RESULT 1
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHRIFFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESSES:
; ADDRESSER: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 936-9300
; TELEFAX: (703) 683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-Flg

US-08-232-463-14

Query Match 8.3%; Score 49.2; DB 2; Length 7218;
Best Local Similarity 5.8%; Pred. No. 0.0065;
Matches 24; Conservative 215; Mismatches 173; Indels 0; Gaps 0;

QY 130 AGCTGGGCCCCGAGCAGCCAGCTCCGAGGAGATGCTTCATCATTAACACCTCCAA 189
DB 1036 AGCTGGCTGAGAGTGAAGGAGCTTGCGATTTTTTTTTTTTTTTTTTTTTTTTTT 1095
QY 190 CAGGCTGAGGCTGAGCTGCGCTGCGCCGCTGCGCCGAGCCGCTTC 249
DB 1096 TT 1155
QY 250 CTGGCGGAGAGATTCTCTCCCTGACGACCATGCTCTATCTCAGGAGCTGAC 309
DB 1156 TT 1215
QY 310 ACCTTCATGATGAGTGAAGCCCTCAGATTCAGTACCTCCCTGGCTCAGAT 369
DB 1216 TT 1275
QY 370 GAAGGCAACCCGAGCTGATGATCTTTAGAAAGCTGAGCTCCGGACTTGGAG 429
DB 1276 TT 1335
QY 430 GACTCTGAGCTGATGATCTTTCTGACATTTGACACATCTCGGTAGAAAAGAGCT 489
DB 1336 TT 1395
QY 490 GCACGGGCCCCGAGCAGCTCTCTCAACCTCTCTGCCCCAGTTCTT 541
DB 1396 TT 1447

RESULT 2

US-10-076-069-1
Sequence 1, Application US/10076069
Patent No. 6872812

GENERAL INFORMATION:
APPLICANT: JURBEC, ROLAND
APPLICANT: NACHTMAN, RONALD
TITLE OF INVENTION: HEPB, A NOVEL GENE WITH A ROLE IN HEMATOPOIETIC AND NEURAL DEVELOPMENT
FILE REFERENCE: 39532-176599
CURRENT APPLICATION NUMBER: US/10/076,069
CURRENT FILING DATE: 2002-02-15
PRIOR APPLICATION NUMBER: US 60/268,923
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PatentIn version 3.0
SEQ ID NO 1
LENGTH: 2082
TYPE: DNA
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: CDS
LOCATION: (191)..(901)
US-10-076-069-1

Query Match 8.0%; Score 47.4; DB 3; Length 2082;
Best Local Similarity 5.0%; Pred. No. 0.013;
Matches 115; Conservative 0; Mismatches 91; Indels 3; Gaps 1;

QY 8 GAGGCTTGAAGAGAAACCTGATTGTTGAAGAGAGAGAGAGTGGAGTGAAGTC 67
DB 201 GAGGCTTGAAGAGAAATATGTTGACCAAGAGAGAGAGTGAAGGTTTGG--GCATG 257
QY 68 CAGCAGGCTTGAAGTCAAGCAGAGCCTGCTCGCATCTCCCTAGCAAAAGTCCAGC 127
DB 258 TCCCTTTCTATAGCTCGAGGAGCAGTCACTCTGACATGTCCTGTCAGAGCTCCAGC 317
QY 128 GAGGCTGGGCCCCGAGCAGCCTTCGAGGAGCATGCTCATCAACACCTTCC 187

DB 318 TCTGTCAATGCTAGTGAAGCCCAATCTGCGCGCTGCTCTCATGCGCAACAGTCC 377
QY 188 AACAGCTGAGGCTGACCTTGCTGCT 216
DB 378 GCGAATTCAGAGAGAAATGAGCAGGAT 406

RESULT 3

US-09-949-016-12197/c
Sequence 12197, Application US/09949016
Patent No. 6812339

GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12197
LENGTH: 44952
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(44952)
OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12197

Query Match 7.9%; Score 46.4; DB 3; Length 44952;
Best Local Similarity 50.4%; Pred. No. 0.059;
Matches 139; Conservative 0; Mismatches 136; Indels 1; Gaps 1;

QY 45 GAGGAGAGTGGAGTGAAGTCCAGAGGCTTTCAGAGCTACCAAGACCCCTGCTCG 104
DB 43405 GAGGAGAGGAGAGAGGGGCGCCAGCAGCTTCCCATCTTCCCGGACCAACCAAG 43346
QY 105 CATCTCCCTAGACAAAGTCCAGCGAGCTGGGCCCCGAGACCCAGCTTCCGAGGCA 164
DB 43345 CCCCCCATCAACCCCTGCCACCCAGCTGGGGCCCCCATCATCTGCTGCACCTGGCCA 43286
QY 165 TGTCTCATCTAATACACCTT-CCAAAGCTGCAAGCTGCACTTGCTGGCTCCGGGCC 223
DB 43285 GGCCTTCATCATATCCCGCCGCGCCAGAGCTGGGGCCCCCAGCATCTGCTGGCTGGGGC 43226
QY 224 CTGCGCTGCCCCCGAGGCCCCCTTCTCTGGGAGAGAGATTTCTCCCTGTCAGCCACCA 283
DB 43225 CTGACTTACTGTATATGTTCTTCCAGAGTGGGGGCTCCACATGTATATCCCTAACCTTCT 43166
QY 284 TTGGCTTATCTCTCAGGAGCTGACACCTTCATG 319
DB 43165 TCCCTCTCGGCTCAGAGATCAGAAACCTCCCAAG 43130

RESULT 4

US-09-949-016-17583/c
Sequence 17583, Application US/09949016
Patent No. 6812339

GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755


```

: OTHER INFORMATION: exon 14
: NAME/KEY: intron
: LOCATION: (46255)..(47035)
: OTHER INFORMATION: intron 14
: NAME/KEY: exon
: LOCATION: (47036)..(47173)
: OTHER INFORMATION: exon 15
: NAME/KEY: intron
: LOCATION: (47174)..(47709)
: OTHER INFORMATION: intron 15
: NAME/KEY: exon
: LOCATION: (47710)..(50544)
: OTHER INFORMATION: exon 16
:
: US-09-733-294A-30

```

Query Match	7.9%;	Score 46.4;	DB 3;	Length 51552;
Best Local Similarity	50.4%;	Pred. No. 0.061;		
Matches 139;	Conservative	0;	Mismatches 136;	Indels 1;
			Gaps	1;
QY	45	GGAGGAGAGGTGGAGGTGAGTCCAGAGGCTCTTCAGAGCTACCAAGACCTGTGTCCG	104	
Db	48820	GGAGGAGAGGGAGAGGGGGGCCAGGCCAGGCTTCCCATTTTCCCGCCAGCCAGACAG	48761	
QY	105	CATCTCCCTAGACAAAGTCAGAGCGCAGCCTTGCGCCCGAGACCCAGCCTCCGACAGCA	164	
Db	48760	CCCCCCCCATCACCCCTGCACCCCACTGAGGGGCCCATCATCTCCGACACCTGGACA	48701	
QY	165	TGTCCTCATCCATTAACACCTT - CCAAGACGTCGAGGCTGCACTTCCGTGCTCCGCCC	223	
Db	48700	GGCCCTCATCATATCCCGCGGCCCAAGCTGGGGCCCCAGCATCCGTGCTCCCTTGGGGC	48641	
QY	224	CTGECCTGGCCCCCGAGACCCCTCTTCTCTGGGCGAGAGAGATTCTCTGTCAAGGACCA	283	
Db	48640	CTGAGACTTACTGTTATGTCTTTCAGGGGTGGGGGTCCCACTGTCTATCCCTTACTCTCT	48581	
QY	284	TTGGCTCATCTCAGGGAGCTGACACTCCANG	319	
Db	48580	TCCCTCTCTGCTCAACGATCAAGAACTCTCCAGG	48545	

```

1 RESULT 6
2 US-10-000-489-71
3 Sequence 71, Application US/10000489
4 Patent No. 6794363
5 GENERAL INFORMATION:
6 APPLICANT: Benjamin, Stephane
7 APPLICANT: Tanaka, Hitoki
8 TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
9 FILE REFERENCE: 91.US6.DIV
10 CURRENT APPLICATION NUMBER: US/10/000,489
11 PRIOR FILING DATE: 2001-11-14
12 PRIOR APPLICATION NUMBER: US 09/924,340
13 PRIOR FILING DATE: 2001-08-06
14 PRIOR APPLICATION NUMBER: PCT/IB01/01715
15 PRIOR FILING DATE: 2001-08-06
16 PRIOR APPLICATION NUMBER: US 60/305,456
17 PRIOR FILING DATE: 2001-07-13
18 PRIOR APPLICATION NUMBER: US 60/302,277
19 PRIOR FILING DATE: 2001-06-29
20 PRIOR APPLICATION NUMBER: US 60/298,698
21 PRIOR FILING DATE: 2001-06-15
22 PRIOR APPLICATION NUMBER: US 60/293,574
23 PRIOR FILING DATE: 2001-05-25
24 NUMBER OF SEQ ID NOS: 112
25 SOFTWARE: JPatent
26 SEQ ID NO 71
27 LENGTH: 1603
28 TYPE: DNA
29 ORGANISM: Homo sapiens
30 FEATURE:
31 NAME/KEY: 5'UTR
32 LOCATION: 1..7
33 NAME/KEY: CDS
34

```

```

? LOCATION: 8..763
? NAME/KEY: 3'UTR
? LOCATION: 764..1603
? NAME/KEY: polyA_signal
? LOCATION: 1562..1567
? NAME/KEY: polyA_site
? LOCATION: 1588..1603
? US-10-000-489-71

```

Query Match	7.2%	Score 42.6;	DB 3;	Length 1603;
Best Local Similarity	54.0%;	Pred. No. 0.22;		
Matches	87;	Conservative	0;	Mismatches 74;
				Indels 0;
				Gaps 0.

Qy	131	GCTGGGCCCCGAGACACCCAGCCTCCGAGAGATGTCATCATCAATAACCTCCAC	130
Db	870	GACAGGAACCCGAGCCATCCCAATCCCCGCGCCGCTCTCCACCCGTGCTTCCC	929
Qy	191	AGTGAAGCTACACTTGCTGGCTTCCGCGCCCTGACCTTACCCCGAGACCCCTCTTC	250
Db	930	CGCTCAACCCCTCACTTACCTCGGCCCCCGGCCACCACTATGGGCCCCGGGCTGTAT	989
Qy	251	TGGGAGAGAGATTTTCCTCTGTACAGCACAATTTGGCTC	251
Db	990	TGTTGGCTGGGCTCGGTGGGCGCTGTCTCTCTCGGCTCT	1030

```

RESULT 7
US-09-249-585A-2/c
? Sequence 2, Application US/09249585A
? Patent No. 6417002
? GENERAL INFORMATION:
? APPLICANT: Horlick, Robert
? TITLE OF INVENTION: METHOD FOR MAINTENANCE AND SELECTION OF EPISOMES
? FILE REFERENCE: 0867/0D905
? CURRENT APPLICATION NUMBER: US/09/249,585A
? CURRENT FILING DATE: 1999-02-11
? NUMBER OF SEQ ID NOS: 18
? SOFTWARE: PatentIn version 3.0
? SEQ ID NO 2
? LENGTH: 1926
? TYPE: DNA
? ORGANISM: Epstein Barr Virus
? FEATURE:
? NAME/KEY: CDS
? LOCATION: (1)..(1926)
? OTHER INFORMATION: coding strand of EBNA-1 DNA
US-09-249-585A-2

```

Query Match 6.9%; Score 41; DB 3; Length 1926

Best Local Similarity 51.4%; Pred. No. 0.59;
Matches 95; Conservative 0; Mismatches 90; Indels 0; Gaps 0

QY	67	CCACAGGACCTTGCAGAGTACACAGCAACCCCTGCTCCGACATCTCCCTGAGCAAAATCCAG	126
Db	1063	CCCGGCGGCTTCCACTACTCTCTCGACCCCGGGCTTCACTACTCTCTGACCCCGGGCTTC	1000
QY	127	CGCAGCCTTGAGCCCCCGAGACACCAGCTCCGCGAGGCAATGCTTCATCCATTAACCCCTC	186
Db	1003	CACATGCCCTCTGAGACCCCGGCGCTGCACCTCTCGTCTCTGACCCCTCTGCTGACCCCTTC	944
QY	187	CAACAGCTGCAAGGCTGACACTTTCGCTGAGCTCCGAGCCCTTGCCCTTACCCCGGAGGCCCTTC	246
Db	943	CTTCTGCTCTCGGCCCTCTCTGCCCCCTCTCTGCTCTCGAGCCCTTCCTGCTCTCTG	884
QY	247	TTTCCT	251
Db	883	CCCCCT	879

RESULT 8
US-09-410-399-3/c
; Sequence 3, Application US/09410399
; Patent No. 6482587


```

; GENERAL INFORMATION:
; APPLICANT: Robertson, Erle S.
; APPLICANT: Colter, Murray A.
; TITLE OF INVENTION: Methods to Inhibit or Enhance the Binding of Viral DNA
; FILE REFERENCE: to Genomic Host DNA
; FILE REFERENCE: UM-03778
; CURRENT APPLICATION NUMBER: US/09/410,399
; CURRENT FILING DATE: 1999-10-01
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1926
; TYPE: DNA
; ORGANISM: Epstein-Barr virus
US-09-410-399-3

Query Match
Best Local Similarity 6.9%; Score 41; DB 3; Length 1926;
Matches 95; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY 67 CCAGCAGGCTTCAAGCTACCAAGCCTGCTCCGATCTCCCTAGCAAAAGTCAG 126
DB 1063 CCCGCGGCTCTCACTACTCTCTCCAGCCCGGCTCACTACTCTCCAGCCCGGCTC 1004
QY 127 CGCAGCTGGGCCCCCGAGCACCAGCCTCCGAGGATGCTCTCATTCATAACACCTTC 186
DB 1003 CACTGCTCTCTGAGCCCGGCTCCAGCTCTGCTCTGCCCCCTCTGCTGCCCCCTC 944
QY 187 CAACAGCTGCAAGCTGCACTTGGCTGCTCCGCCCCCTGCTGCCCCCGAGCCCTTC 246
DB 943 CTCCTGCTCTCTGCTCTCTCTGCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTG 884
QY 247 TTCCT 251
DB 883 CCCCT 879

RESULT 9
US-09-050-863-2/c
; Sequence 2, Application US/09050863
; GENERAL INFORMATION:
; APPLICANT: Lao, Ying
; APPLICANT: Hiang, Betty
; APPLICANT: Payan, Don
; TITLE OF INVENTION: Mammalian Protein Interaction Cloning
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Flehr, Hohbach, Test, Albritton & Herbert
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/050,863
; FILING DATE: 30-MAR-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Silva, Robin M.
; REGISTRATION NUMBER: 38,304
; REFERENCE/DOCKET NUMBER: A-65638/DJB/RMS
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 949-8711
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:

```

```

; LENGTH: 2580 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA
US-09-050-863-2

Query Match
Best Local Similarity 6.9%; Score 41; DB 3; Length 2580;
Matches 95; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY 67 CCAGCAGGCTTCAAGCTACCAAGCCTGCTCCGATCTCCCTAGCAAAAGTCAG 126
DB 1446 CCCGCGGCTCTCACTACTCTCTCCAGCCCGGCTCACTACTCTCCAGCCCGGCTC 1387
QY 127 CGCAGCTGGGCCCCCGAGCACCAGCCTCCGAGGATGCTCTCATTCATAACACCTTC 186
DB 1386 CACTGCTCTCTGAGCCCGGCTCCAGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTG 1327
QY 187 CAACAGCTGCAAGCTGCACTTGGCTGCTCCGCCCCCTGCTGCCCCCGAGCCCTTC 246
DB 1326 CTCCTGCTCTCTGCTCTCTCTGCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTG 1267
QY 247 TTCCT 251
DB 1266 CCCCT 1262

RESULT 10
US-09-359-081-2/c
; Sequence 2, Application US/09359081
; Patent No. 6316223
; GENERAL INFORMATION:
; APPLICANT: Lao, Ying
; APPLICANT: Hiang, Betty
; APPLICANT: Payan, Don
; TITLE OF INVENTION: Mammalian Protein Interaction Cloning
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Flehr, Hohbach, Test, Albritton & Herbert
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/359,081
; FILING DATE: 22-Jul-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/050,863
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Silva, Robin M.
; REGISTRATION NUMBER: 38,304
; REFERENCE/DOCKET NUMBER: A-65638/DJB/RMS
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 949-8711
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2580 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:

```

US-09-359-081-2

Query Match 6.9%; Score 41; DB 3; Length 2580;
Best Local Similarity 51.4%; Pred. No. 0.65;
Matches 95; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY 67 CCAGCAGGCTTCAAGCTACCAAGCAGCCTGCTCCGATCTCCCTGACAAAGTCCAG 126
DB 1446 CCGGCGGCTCCCACTACCTCTCGACCCCGGCTCCCACTACCTCTCGACCCCGGCTTC 1387
QY 127 CGCAGCCTGGGCGCCCGAGCAGCCGAGCCTCGAGAGATGCTGATCATTAACACCTTC 186
DB 1386 CACTGCTCTCTGACCCCGGCTCCCACTCTGCTCTGCCCCCTCTGCTCTGCCCCCTC 1327
QY 187 CAACAGCTGACAGGCTGCACTTCTGCTGCTCCGAGCCTCGAGCCTGCCCCCGAGCCCTTC 246
DB 1326 CTCCTGCTCTGCCCCCTCTGCTCTGCTCTGCTCTGCCCCCTCTGCTCTGCTCTG 1267
QY 247 TTCTT 251
DB 1266 CCCCC 1262

RESULT 11

US-09-130-114-1

/ Sequence 1, Application US/09130114
/ Patent No. 5976807
/ GENERAL INFORMATION:
/ APPLICANT: Horlick, Robert A.
/ APPLICANT: Dama, Bassem B.
/ APPLICANT: Robbins, Alan K.
/ TITLE OF INVENTION: Eukaryotic Cells Stably Expressing Genes
/ FILE REFERENCE: 0867/ID903US1
/ CURRENT APPLICATION NUMBER: US/09/130,114
/ CURRENT FILING DATE: 1998-08-06
/ NUMBER OF SEQ ID NOS: 36
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 1
/ LENGTH: 5452
/ TYPE: DNA
/ ORGANISM: VEBNA
US-09-130-114-1

Query Match 6.9%; Score 41; DB 2; Length 5452;

Best Local Similarity 51.4%; Pred. No. 0.8;
Matches 95; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY 67 CCAGCAGGCTTCAAGCTACCAAGCAGCCTGCTCCGATCTCCCTGACAAAGTCCAG 126
DB 1359 CCGGCGGCTCCCACTACCTCTCGACCCCGGCTCCCACTACCTCTCGACCCCGGCTTC 1418
QY 127 CGCAGCCTGGGCGCCCGAGCAGCCGAGCCTCGAGAGATGCTGATCATTAACACCTTC 186
DB 1419 CACTGCTCTCTGACCCCGGCTCCCACTCTGCTCTGCCCCCTCTGCTCTGCCCCCTC 1478
QY 187 CAACAGCTGACAGGCTGCACTTCTGCTGCTCCGAGCCTCGAGCCTGCCCCCGAGCCCTTC 246
DB 1479 CTCCTGCTCTGCCCCCTCTGCTCTGCTCTGCTCTGCCCCCTCTGCTCTGCTCTG 1538
QY 247 TTCTT 251
DB 1539 CCCCC 1543

RESULT 12

US-09-647-344A-14

/ Sequence 14, Application US/09647344A
/ Patent No. 6586180
/ GENERAL INFORMATION:
/ APPLICANT: Ruffner, Duane E.
/ APPLICANT: Pierce, Michael L.
/ APPLICANT: Chen, Zhidong

/ TITLE OF INVENTION: Directed Antisense Libraries

/ FILE REFERENCE: 16678.PCT-US

/ CURRENT APPLICATION NUMBER: US/09/647,344A

/ CURRENT FILING DATE: 2000-12-04

/ PRIOR APPLICATION NUMBER: PCT/US99/06742

/ PRIOR FILING DATE: 1999-03-28

/ NUMBER OF SEQ ID NOS: 50

/ SEQ ID NO 14

/ LENGTH: 8705

/ TYPE: DNA

/ ORGANISM: Artificial Sequence

/ FEATURE:

/ OTHER INFORMATION: pshuttle

US-09-647-344A-14

Query Match 6.9%; Score 41; DB 3; Length 8705;
Best Local Similarity 51.4%; Pred. No. 0.92;
Matches 95; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY 67 CCAGCAGGCTTCAAGCTACCAAGCAGCCTGCTCCGATCTCCCTGACAAAGTCCAG 126
DB 7224 CCGGCGGCTCCCACTACCTCTCGACCCCGGCTCCCACTACCTCTCGACCCCGGCTTC 7283
QY 127 CGCAGCCTGGGCGCCCGAGCAGCCGAGCCTCGAGAGATGCTGATCATTAACACCTTC 186
DB 7284 CACTGCTCTCTGACCCCGGCTCCCACTCTGCTCTGCCCCCTCTGCTCTGCCCCCTC 7343
QY 187 CAACAGCTGACAGGCTGCACTTCTGCTGCTCCGAGCCTCGAGCCTGCCCCCGAGCCCTTC 246
DB 7344 CTCCTGCTCTGCCCCCTCTGCTCTGCTCTGCTCTGCCCCCTCTGCTCTGCTCTG 7403
QY 247 TTCTT 251
DB 7404 CCCCC 7408

RESULT 13

US-08-910-647-1/c

/ Sequence 1, Application US/08910647
/ Patent No. 6251433
/ GENERAL INFORMATION:

/ APPLICANT: Zuckermann et al.

/ TITLE OF INVENTION: Compositions and Methods for

/ NUMBER OF SEQUENCES: 4

/ CORRESPONDENCE ADDRESS:

/ ADDRESS: Chiron Corporation

/ STREET: 4560 Horton Street

/ CITY: Emeryville

/ STATE: California

/ COUNTRY: U.S.A.

/ ZIP: 94608-2916

/ COMPUTER READABLE FORM:

/ MEDIUM TYPE: Floppy disk

/ OPERATING SYSTEM: PC-DOS/MS-DOS

/ SOFTWARE: Patent in Release #1.0, Version #1.30

/ CURRENT APPLICATION DATA:

/ APPLICATION NUMBER: US/08/910,647

/ FILING DATE:

/ CLASSIFICATION: 514

/ ATTORNEY/AGENT INFORMATION:

/ NAME: Fujita, Sharon M.

/ REGISTRATION NUMBER: 38,459

/ TELECOMMUNICATION INFORMATION:

/ TELEPHONE: (510) 923-2706

/ TELEFAX: (510) 655-3542

/ INFORMATION FOR SEQ ID NO: 1:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 9600 base pairs

/ TYPE: nucleic acid

/ STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-910-647-1

Query Match
Best Local Similarity 51.4%; Score 41; DB 3; Length 9600;
Matches 95; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY 67 CCAGCAGGCTTCAAGCTACCAAGAGCCCTGCTCCGATCTCCCTAGACAAAGTCCAG 126
DB 1492 CCCGGGGGCTCCACTACTCTCTCCAGACCCGGGCTCTCACTACTCTGACCCCGGCTTC 1433
QY 127 CGCAGCTGGGCCCCCGAGACCCAGCCTCCGAGGAGTCTCTATCCATAACACCTTC 186
DB 1432 CACTGCTCTCTGACCCCGGCTCCAGCTCTGCTCTGCCCCCTCTGCTCTGCCCCCTTC 1373
QY 187 CAACAGCTGACAGCTGCACTTGCGCTGCTCCCGGCTCCGCTGCCCCCGAGCCCCCTTC 246
DB 1372 CTCTGCTCTCTGCCCCCTCTGCTCTGCTCTGCTCTGCCCCCTCTGCTCTGCTCTG 1313
QY 247 TTCTCT 251
DB 1312 CCCCT 1308

RESULT 14

US-09-620-925-1/c
Sequence 1, Application US/09620925
Patent No. 6468986

GENERAL INFORMATION:

APPLICANT: Zuckermann et al.
TITLE OF INVENTION: Compositions and Methods for
Polynucleotide Delivery

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: Chiron Corporation
STREET: 4560 Horton Street
CITY: Emeryville
STATE: California
COUNTRY: U.S.A.
ZIP: 94608-2916

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/620,925
FILING DATE: 21-Jul-2000
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/910,647
FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Fujita, Sharon M., 459
REGISTRATION NUMBER: 38,459
REFERENCE/DOCKET NUMBER: 1218.002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 923-2706
TELEFAX: (510) 655-3542

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
LENGTH: 9600 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-620-925-1

Query Match
Best Local Similarity 51.4%; Score 41; DB 3; Length 9600;
Matches 95; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY 67 CCAGCAGGCTTCAAGCTACCAAGAGCCCTGCTCCGATCTCCCTAGACAAAGTCCAG 126
DB 1492 CCCGGGGGCTCCACTACTCTCTCCAGACCCGGGCTCTCACTACTCTGACCCCGGCTTC 1433
QY 127 CGCAGCTGGGCCCCCGAGACCCAGCCTCCGAGGAGTCTCTATCCATAACACCTTC 186
DB 1432 CACTGCTCTCTGACCCCGGCTCCAGCTCTGCTCTGCCCCCTCTGCTCTGCCCCCTTC 1373
QY 187 CAACAGCTGACAGCTGCACTTGCGCTGCTCCCGGCTCCGCTGCCCCCGAGCCCCCTTC 246
DB 1372 CTCTGCTCTCTGCCCCCTCTGCTCTGCTCTGCTCTGCCCCCTCTGCTCTGCTCTG 1313
QY 247 TTCTCT 251
DB 1312 CCCCT 1308

RESULT 15

US-07-884-811-15/c
Sequence 15, Application US/07884811
Patent No. 5316921

GENERAL INFORMATION:

APPLICANT: Godowski, Paul J, Lohker, Nathalie A, Mark, Melanie R.
TITLE OF INVENTION: SINGLE-CHAIN HEPATOCTE GROWTH FACTOR VARIANTS
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/884,811
FILING DATE: 19920518
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER:

ATTORNEY/AGENT INFORMATION:

NAME: Dreger, Ginger R.
REGISTRATION NUMBER: 33,055
REFERENCE/DOCKET NUMBER: 755.1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-3216
TELEFAX: 415/952-9881
TELEX: 910/371-7168

INFORMATION FOR SEQ ID NO: 15:

SEQUENCE CHARACTERISTICS:
LENGTH: 10596 bases
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
US-07-884-811-15

Query Match
Best Local Similarity 51.4%; Score 41; DB 2; Length 10596;
Matches 95; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY 67 CCAGCAGGCTTCAAGCTACCAAGAGCCCTGCTCCGATCTCCCTAGACAAAGTCCAG 126
DB 3026 CCCGGGGGCTCCACTACTCTCTCCAGACCCGGGCTCTCACTACTCTGACCCCGGCTTC 2967
QY 127 CGCAGCTGGGCCCCCGAGACCCAGCCTCCGAGGAGTCTCTATCCATAACACCTTC 186
DB 2966 CACTGCTCTCTGACCCCGGCTCCAGCTCTGCTCTGCCCCCTCTGCTCTGCTCTG 2907

QY	187	CAACGCTGAGGCGTCACTTGCCTGGGTCCCGCCCTGACCTGCCCCGAGCCCTG	246
Db	2906	CTCTGCTCTCTGCGCCCTCTGCGCCCTCTGCTCTGCGCCCTCTGCGCCCTCTGCTCTG	2844
QY	247	TTGCT	251
Db	2846	CCGCT	2842

Search completed: February 6, 2006, 11:23:17
Job time : 178 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 6, 2006, 11:31:10 ; Search time 741 Seconds
(without alignments)
6595.415 Million cell updates/sec

Title: US-10-069-386a-1

Perfect score: 591
Sequence: 1 atgagagagagcttgaagag.....aaatcatctcgggtcctaa 591

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA Main:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	578.4	97.9	598	US-10-029-386-24898	Sequence 24898, A
2	576.2	97.5	1596	US-10-106-698-1239	Sequence 1239, Ap
3	473	80.0	506	US-09-867-701-6064	Sequence 6064, Ap
4	186.6	31.6	205	US-09-867-701-6239	Sequence 6239, Ap
5	166	28.1	526	US-10-029-386-11178	Sequence 11178, A
6	47.4	8.0	2082	US-10-076-065-1	Sequence 1, Appl
7	46.4	7.9	3179	US-10-840-455-20	Sequence 20, Appl
8	46.4	7.9	25138	US-10-840-455-44	Sequence 44, Appl
9	46.4	7.9	51552	US-09-733-294A-30	Sequence 30, Appl
10	44.8	7.6	440	US-10-184-644-202	Sequence 202, App
11	44.8	7.6	440	US-10-184-634-202	Sequence 202, App
12	44.8	7.6	440	US-10-063-685-52	Sequence 52, Appl
13	44.8	7.6	594	US-10-123-155-10	Sequence 10, Appl
14	44.8	7.6	594	US-10-146-731-10	Sequence 10, Appl
15	44.8	7.6	594	US-10-140-472-10	Sequence 10, Appl
16	44.8	7.6	594	US-10-141-761-10	Sequence 10, Appl
17	44.8	7.6	594	US-10-142-885-10	Sequence 10, Appl
18	44.8	7.6	594	US-10-158-790-10	Sequence 10, Appl
19	44.8	7.6	594	US-10-137-871-10	Sequence 10, Appl
20	44.8	7.6	594	US-10-140-923-10	Sequence 10, Appl
21	44.8	7.6	594	US-10-141-756-10	Sequence 10, Appl
22	44.8	7.6	594	US-10-141-759-10	Sequence 10, Appl
23	44.8	7.6	594	US-10-140-805-10	Sequence 10, Appl

C	24	44.8	7.6	594	6	US-10-140-864-10	Sequence 10, Appl
	25	44.8	7.6	2748	7	US-10-437-963-53188	Sequence 53188, A
	26	42.6	7.2	1603	3	US-09-992-600A-71	Sequence 71, Appl
	27	42.6	7.2	1603	3	US-09-924-340-71	Sequence 71, Appl
	28	42.6	7.2	1603	3	US-09-992-095B-71	Sequence 71, Appl
	29	42.6	7.2	1603	3	US-09-999-570-71	Sequence 71, Appl
	30	42.6	7.2	1603	5	US-10-000-489-71	Sequence 71, Appl
	31	42.6	7.2	1603	5	US-10-000-988-71	Sequence 71, Appl
	32	42.6	7.2	1603	6	US-10-154-678-71	Sequence 71, Appl
	33	42.6	7.2	1603	6	US-10-001-142-71	Sequence 71, Appl
	34	42.6	7.2	1603	8	US-10-838-854-71	Sequence 71, Appl
	35	42.6	7.2	511	6	US-10-029-386-24981	Sequence 24981, A
	36	42.4	7.2	600	8	US-10-363-345A-39087	Sequence 39087, A
	37	42.4	7.2	600	8	US-10-363-345A-39088	Sequence 39088, A
	38	42.4	7.2	600	9	US-10-363-483A-39087	Sequence 39087, A
	39	42.4	7.2	600	9	US-10-363-483A-39088	Sequence 39088, A
	40	42.4	7.2	1650	7	US-10-437-963-27296	Sequence 27296, A
	41	42	7.1	672	7	US-10-437-963-69519	Sequence 69519, A
	42	42	7.1	2580	6	US-10-156-761-3895	Sequence 3895, Ap
	43	42	7.1	9025608	6	US-10-156-761-1	Sequence 1, Appl
	44	41.8	7.1	32329	7	US-10-374-903A-1	Sequence 1, Appl
	45	41.6	7.0	940	7	US-10-437-963-40850	Sequence 40850, A

ALIGNMENTS

```
RESULT 1
US-10-029-386-24898
; Sequence 24898, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn. Sharron G.
; APPLICANT: Rank, David R.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: AEWICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 24898
; LENGTH: 598
; TYPER: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC010271.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.64
; OTHER INFORMATION: EST HUMAN HIT: BG720189.1, EVALUR 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: 070622, EVALUR 2.30e-01
; OTHER INFORMATION: NT HIT: g116178011, EVALUR 0.00e+00
US-10-029-386-24898

Query Match      97.9%; Score 578.4; DB 6; Length 598;
Best Local Similarity 97.7%; Pred. No. 4.1e-161;
Matches 590; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY      1 ATGAGGAGAGGCTTGAAGAGAAACCTCTGATTGGAAGAGAGAGAGAGGAGAG 60
Db      7 ATGGTGGAGAGGCTTGAAGAGAAACCTCTGATTGGAAGAGAGAGAGAGGAGAG 66

QY      61 TGGAGTCCAGCAGGCTTCAAGGCTACAGCAAGGCTCTCCGATCTCCCTTAACAA 120
Db      67 TGGAGTCCAGCAGGCTTCAAGGCTACAGCAAGGCTCTCCGATCTCCCTTAACAA 126

QY      121 GTTCAGGAGAGGCTGGGCCCCGAGACACCACTCCGAGGAGATGCTTCAATAC 180
Db      127 GTTCAGGAGAGGCTGGGCCCCGAGACACCACTCCGAGGAGATGCTTCAATAC 186

QY      181 ACCCTTCAACAGCTGCAAGCTGCACTGCTGCTCCGCTCCGCTCCGCTCCGCTCCG 239
Db      187 ACCCTTCAACAGCTGCAAGCTGCACTGCTGCTCCGCTCCGCTCCGCTCCGCTCCG 246
```

QY 240 GCGGCTCTTCTGCGGCGAGAGATTTCCTCTGTCAGCCACCATTTGGCTTATCTCAG 299
DB 247 GCGGCTCTTCTGCGGCGAGAGATTTCCTCTGTCAGCCACCATTTGGCTTATCTCAG 306
QY 300 GGAAGTGGACACCTCCATGATGATGGGACTGAGCCCCCTCAGAAATCCAGGACTCCCTTGG 359
DB 307 GGAAGTGGACACCTCCATGATGATGGGACTGAGCCCCCTCAGAAATCCAGGACTCCCTTGG 366
QY 360 CCTCCAGATGAATGAGCCACCCAGCCTGATCCAGTCTTCTTGAAGCTGAGCTCCCG 419
DB 367 CCTCCAGATGAATGAGCCACCCAGCCTGATCCAGTCTTCTTGAAGCTGAGCTCCCG 426
QY 420 GTACTTGGGGGACTCTGAGCTGATGACTTCTTCTGACATTTGACACATCTGCCGTAGA 479
DB 427 GTACTTGGGGGACTCTGAGCTGATGACTTCTTCTGACATTTGACACATCTGCCGTAGA 486
QY 480 AAAAGAGCTGACCGGGGCCCCACGAGAGCTCCCTCAGAACCTCTGAGCCCCAGGTTTC 539
DB 487 AAAAGAGCTGACCGGGGCCCCACGAGAGCTCCCTCAGAACCTCTGAGCCCCAGGTTTC 546
QY 540 TTGGAGTGAATGAATGATGATCATCATGAGAAATCATTTGGGGTCTTAA 591
DB 547 TTGGAGTGAATGAATGATGATCATCATGAGAAATCATTTGGGGTCTTAA 598

RESULT 2

US-10-106-698-1239
; Sequence 1239, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; PRIOR FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 1239
; LENGTH: 1596
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURES:
; NAME/KEY: misc_feature
; LOCATION: (57)..(57)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-106-698-1239

Query Match 97.5%; Score 576.2; DB 5; Length 1596;
Best Local Similarity 99.3%; Pred. No. 2,1e-160;
Matches 587; Conservative 2; Mismatches 0; Indels 2; Gaps 1;

QY 1 ATGGAGGAGGCTTTAAAGAGAAACACTGTGATTTGGAGAGAGAGAGAGAGGTTGGAG 60
DB 345 ATGGAGGAGGCTTTAAAGAGAAACACTGTGATTTGGAGAGAGAGAGAGAGGTTGGAG 404
QY 61 TGAAGTCCAGAGGCTTTCAGAGCTACAGAGAGGCTCTCGCATCTCCCTAGACAA 120
DB 405 TGAAGTCCAGAGGCTTTCAGAGCTACAGAGAGGCTCTCGCATCTCCCTAGACAA 464
QY 121 GTCCAGCGAGGCTTGGGCCCCGAGAGACCCAGGCTTCGAGAGGATGTCTCATTCATAAC 180
DB 465 GTCCAGCGG--CTGGGCCCCGAGAGACCCAGGCTTCGAGAGGATGTCTCATTCATAAC 522
QY 181 ACCCTTCAACAGCTGAGGCTGACATTTGGCTGGCTCCGCGCCCTGGCCCCCGAG 240
DB 523 ACCCTTCAACAGCTGAGGCTGACATTTGGCTGGCTCCGCGCCCTGGCCCCCGAG 582

QY 241 CCGCTTCTTCTGGGCGAGAGATTTCCTCTGTCAGCCACCATTTGGCTTATCTCAG 300
DB 583 CCGCTTCTTCTGGGCGAGAGATTTCCTCTGTCAGCCACCATTTGGCTTATCTCAG 642
QY 301 GAGCTGGACACTCCATGATGATGGGACTGAGCCCCCTCAGAAATCCAGGACTCCCTTGGC 360
DB 643 GAGCTGGACACTCCATGATGATGGGACTGAGCCCCCTCAGAAATCCAGGACTCCCTTGGC 702
QY 361 CTCAGAAATGAATGAGCCACCCAGCCTGATCCAGTCTTCTTGAAGCTTGAAGTCCCG 420
DB 703 CTCAGAAATGAATGAGCCACCCAGCCTGATCCAGTCTTCTTGAAGCTTGAAGTCCCG 762
QY 421 TACTTGGGGGACTCTGAGCTGATGACTTCTTCTGACATTTGACACATCTGCCGTAGA 480
DB 763 TACTTGGGGGACTCTGAGCTGATGACTTCTTCTGACATTTGACACATCTGCCGTAGA 822
QY 481 AAGAGAGCTGACCGGGGCCCCACGAGAGCTCCCTCAGAACCTCTTGTGCCCCAGTTCT 540
DB 823 AAGAGAGCTGACCGGGGCCCCACGAGAGCTCCCTCAGAACCTCTTGTGCCCCAGTTCT 882
QY 541 TTGGAGTGAATGAATGATGATCATCATGAGAAATCATTTGGGGTCTTAA 591
DB 883 TTGGAGTGAATGAATGATGATCATCATGAGAAATCATTTGGGGTCTTAA 933

RESULT 3

US-09-867-701-6064
; Sequence 6064, Application US/09867701
; Patent No. US2002013237A1
; GENERAL INFORMATION:
; APPLICANT: Aglate, Paul A.
; APPLICANT: Jones, Robert
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; PRIOR FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6064
; LENGTH: 506
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-867-701-6064

Query Match 80.0%; Score 473; DB 3; Length 506;
Best Local Similarity 98.6%; Pred. No. 6.9e-130;
Matches 488; Conservative 0; Mismatches 5; Indels 2; Gaps 1;

QY 60 GTGAGTCCAGAGGCTTTCAGAGCTACAGAGAGGCTCTCGCATCTCCCTAGACAA 119
DB 14 GCGAGTCCAGAGGCTTTCAGAGCTACAGAGAGGCTCTCGCATCTCCCTAGACAA 73
QY 120 AGTCCAGGCGAGGCTTGGGCCCCGAGAGACCCAGGCTTCGAGAGGATGTCTCATCATTA 179
DB 74 AGTCCAGGCG--CTGGGCCCCGAGAGACCCAGGCTTCGAGAGGATGTCTCATCATTA 131
QY 180 CACCTTCCAGAGGCTGAGGCTGCACTTGAGCTCCGCGCCCTGGCCCCCGAG 239
DB 132 CACCTTCCAGAGGCTGAGAGATGCACTTGAGCTCCGCGCCCTGGCCCCCGAG 191
QY 240 GCGCTTCTTCTGGGCGAGAGATTTCCTCTGTCAGCCACCATTTGGCTTATCTCAG 299
DB 192 GCGCTTCTTCTGGGCGAGAGATTTCCTCTGTCAGCCACCATTTGGCTTATCTCAG 251
QY 300 GGAAGTGGACACTCCATGATGATGGGACTGAGCCCCCTCAGAAATCCAGGACTCCCTTGG 359
DB 252 GGAAGTGGACACTCCATGATGATGGGACTGAGCCCCCTCAGAAATCCAGGACTCCCTTGG 311
QY 360 CCTCCAGAAATGAATGAGCCACCCAGCCTGATCCAGTCTTCTTGAAGCTTGAAGCTCCCG 419

Db 312 CCTCCAGAAATGAGTGCACCCAGCTGATCCAGTCTTCTTAGAAGCTCTGAGCTCCCG 371
QY 420 GTACTTGGGGGAGTCTGAGCTGATGATCTTTCTTGACATTTGACATCTGCGGTAGA 479
Db 372 GTACTTGGGGGAGTCTGAGCTGATGATCTTTCTTGACATTTGACATCTGCGGTAGA 431
QY 480 AAGAGAGCTGACAGGGGCCCCAGAGAGCTCTCTCAACCTTTCTGTGCCCCAGTTTC 539
Db 432 AAGAGAGCTGACAGGGGCCCCAGAGAGCTCTCTCAACCTTTCTGTGCCCCAGTTTC 491
QY 540 TTGGAGTGTGATGA 554
Db 492 TTGGAGTGTGATGA 506

RESULT 4
US-09-867-701-6239
; Sequence 6239, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Aglate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6239
; LENGTH: 205
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-867-701-6239

Query Match 31.6%; Score 186.6; DB 3; Length 205;
Best Local Similarity 97.1%; Pred. No. 4.8e-45;
Matches 201; Conservative 0; Mismatches 4; Indels 2; Gaps 1;
QY 63 GAGTCAGAGAGCTTTCAGAGCTTACAGAGAGAGCTTCTGCGATCTCCCTAGACAAAGT 122
Db 1 GAGTCAGAGAGCTTTCAGAGCTTACAGAGAGAGCTTCTGCGATCTCCCTAGACAAAGT 60
QY 123 CAGAGGAGAGCTTGGGCCCCGAGAGAGAGCTTCTGCGAGAGAGATCTCTATCTATAC 182
Db 61 CAGAGGAG--CTTGGGCCCCGAGAGAGAGAGCTTCTGCGAGAGATCTCTATCTATAC 118
QY 183 CCTCCAGAGCTGACAGGCTGACCTTGCTGGCTCCCGGCTGAGCTGAGCTGAGAGCC 242
Db 119 CCTCCAGAGCTGACAGGCTGACCTTGCTGGCTCCCGGCTGAGCTGAGAGCC 178
QY 243 CCTTCTCTGGGAGAGAGATTTCTC 269
Db 179 CCTTCTCTGGGAGAGAGATTTCTC 205

RESULT 5
US-10-029-386-11178
; Sequence 11178, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: ABOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 11178

; LENGTH: 526
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC010271.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.64
; OTHER INFORMATION: SWISSPROT HIT: Q60963, EVALU 2.20e+00
; OTHER INFORMATION: NT HIT: 9115718683, EVALU 0.00e+00
; OTHER INFORMATION: EST_HUMAN HIT: AA426355.1, EVALU 1.00e-114
US-10-029-386-11178

Query Match 28.1%; Score 166; DB 6; Length 526;
Best Local Similarity 100.0%; Pred. No. 7e-39;
Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 426 GGGGAGCTTGGGCTGAGATGATCTTCTTCTGACATTTGACATCTGCGGTAGAAAGGA 485
Db 1 GGGGAGCTTGGGCTGAGATGATCTTCTTCTGACATTTGACATCTGCGGTAGAAAGGA 60
QY 486 GCTTCAGAGGGGCCCCAGAGAGCTTCTCAACCTTTCTGTGCCCCAGTTCTTGGGA 545
Db 61 GCTTCAGAGGGGCCCCAGAGAGCTTCTCAACCTTTCTGTGCCCCAGTTCTTGGGA 120
QY 546 GTGAATGAATGATGATCATCATGGAATCATTTCTGGGCTCTTA 591
Db 121 GTGAATGAATGATGATCATCATGGAATCATTTCTGGGCTCTTA 166

RESULT 6
US-10-076-069-1
; Sequence 1, Application US/10076069
; Publication No. US2002017214A1
; GENERAL INFORMATION:
; APPLICANT: JURECIC, ROLAND
; APPLICANT: NACHTMAN, RONALD
; TITLE OF INVENTION: HEPP, A NOVEL GENE WITH A ROLE IN HEMATOPOIETIC AND NEURAL DEVELOP
; FILE REFERENCE: 39532-176599
; CURRENT APPLICATION NUMBER: US/10/076,069
; CURRENT FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: US 60/268,923
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 2082
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (191)..(901)
US-10-076-069-1

Query Match 8.0%; Score 47.4; DB 5; Length 2082;
Best Local Similarity 55.0%; Pred. No. 0.0012;
Matches 115; Conservative 0; Mismatches 91; Indels 3; Gaps 1;
QY 8 GAGGCTTGAAGAGAAACATCTGATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 67
Db 201 GAGGCTTGAAGAGAAATATGATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 257
QY 68 CAGCAGAGCTTTCAGAGCTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 127
Db 258 TCCCTTCTATATAGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 317
QY 128 GAGGCTTGGGCCCCGAG 187
Db 318 TCTGTCAATGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 377
QY 188 AACAGTGAAG 216
Db 378 GGCAGATCCAGAT 406

```

RESULT 7
US-10-840-455-20/c
; Sequence 20, Application US/10840455
; Publication NO. US20050032094A1
; GENERAL INFORMATION:
; APPLICANT: Bayer Aktiengesellschaft
; APPLICANT: Hagen, Gustav
; APPLICANT: Wick, Mareesa
; APPLICANT: Zubov, Dmitry
; TITLE OF INVENTION: Regulatory DNA Sequences of the Gene for the Human Catalytic
; TITLE OF INVENTION: Telomerase Subunit, and Their Diagnostic and Therapeutic Use
; FILE REFERENCE: Lea 32 805C1
; CURRENT APPLICATION NUMBER: US/10/840,455
; CURRENT FILING DATE: 2004-05-06
; PRIOR APPLICATION NUMBER: PCT/EP98/08216
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: US 09/582,246
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: DE19757984.1
; PRIOR FILING DATE: 1997-12-24
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 20
; LENGTH: 3179
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-840-455-20

```

Query Match	7.9%	Score 46.4;	DB 8;	Length 3179;
Best Local Similarity	50.4%;	Pred. No. 0.0026;		
Matches 139;	Conservative 0;	Mismatches 136;	Indels 1;	Gaps 1.

Oy	45	GGAGAGAGAGTGGAGTGAAGTCACAGAGGCTTCAGAGGTACACAGAGGCTGTCCG	104
Db	447	GGAGGACAGGCGAGAGGGGGCCAGGCAAGCTTCCCATTTTCCCGGCACCCAGACAG	388
Oy	105	CATCTTCCTTAAGACAAAGTCAGAGGCGAGGCTGTGGGCCCTCCAGACACCGACTCCGACAGCA	164
Db	387	CCCCCCCCATCAACCCCTGCGACACCCCAAGCTGGGGCCCCCATCATCCTGTGCAACCTTGCCCA	328
Oy	165	TGTGCTCATCCATTAACACT--CCAAAGCTGAGAGGTGACTTGTGGCTGGCTCCGACC	223
Db	327	GGCCCTTCATCATCCCGCGCGCCAAAGCTGGGCCCCCAGCATTCCTGTGCTTGGGC	268
Oy	224	CTGCCCTGGCCCCCGAGGCCCTCTTCTCTGGGCGAGAGGATTTCTCCCTGTGACGACCA	283
Db	267	CTGAGACTTACTGTATGTCTTCCAGAGGTGGGGCTCCCACTGTCTATCCCTCACTCTCT	208
Oy	284	TTGGCTCTATCTCTCAAGGAGCTGGAACACTCCATAG	319
Db	207	TCCCTCTCGCTGCAGGATCAGAAACCTCCAG	172

RESULT 8
 US-10-840-455-44/c
 : Sequence 44, Application US/10840455
 : Publication No. US20050032054A1
 :
 : GENERAL INFORMATION:
 :
 : APPLICANT: Bayer Aktiengesellschaft
 : APPLICANT: Hagen, Gustav
 : APPLICANT: WICK, Mareisa
 : APPLICANT: Zubov, Dmitry
 :
 : TITLE OF INVENTION: Regulatory DNA Sequences of the Gene for the Human Catalytic
 : TITLE OF INVENTION: Telomerase Subunit, and Their Diagnostic and Therapeutic Use
 : FILE REFERENCE: Lea 32 805051
 :
 : CURRENT APPLICATION NUMBER: US/10/840,455
 :
 : CURRENT FILING DATE: 2004-05-06
 :
 : PRIOR APPLICATION NUMBER: PCT/EP98/08216
 :
 : PRIOR FILING DATE: 1998-12-22
 :
 : PRIOR APPLICATION NUMBER: US 09/582,246
 :
 : PRIOR FILING DATE: 2000-09-21
 :
 : PRIOR APPLICATION NUMBER: DE19757984.1
 :
 : PRIOR FILING DATE: 1997-12-24

```

? NUMBER OF SEQ ID NOS: 44
? SOFTWARE: PatentIn version 3.2
? SEQ ID NO 44
? LENGTH: 25138
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: misc feature
? LOCATION: (1)..(25138)
? OTHER INFORMATION: n is a, c, g, or t
US-10-840-435-44

```

Query Match	7.9%	Score 46.4	DB 8	Length 25138
Best Local Similarity	50.4%	Pred. No. 0.0035		
Matches 139; Conservative	0	Mismatches 136	Indels 1	Gaps 1

Oy	45	GAGAGAAAGGTGGAGAAGTGCACAGAGGCTTCAGAGATCAGACAAGCCCTGTCCG	104
Db	22406	GGAAGCAAGGAGAGAGGGGCCAAGCAGGCTTCCCATTCTTCCCGCACCCAGACAG	22347
Oy	105	CATCTCCCTTAGAACAAAGTCAGGCGCAGCTTGGGCCCCGAGACACCAGCTTCGACAGCA	164
Db	22346	CCCCCCCCCATCACCCCTGCACACCCACAGCTGAGGGCCCCCATCATCTCTGCACACCTGGCCA	22287
Oy	165	TGTCCTCATCATTAACAACCT - CCAACAGCTGAGAGGTGACTTTGGCTGGCTCCGAGCC	223
Db	22286	GAGCCTTCATCATATCCCGCGCCCAAGCTGAGGCCCCCAAGATCTCTGTGCTTCGGAGC	22222
Oy	224	CTGACCTGACCCCCGAGGCCCTCTTCTCTGAGCAGAGAGATTTCTCTCTGTACAGCCACA	283
Db	22226	CTGGAATTACTGTATGTATCTTTCAGAGGATGAGGAGCTCCCACTGCTCATCCCTACCTCTCT	22167
Oy	284	TTGGCTTATCTCAAGAGAGCTGAAACCTCCANTG	319
Db	22166	TCCCTCTCTGCATCAGATCAAGAAACCTCCAGG	22131

```

1 RESULT 9
2 US-09-733-294A-30/C
3 Sequence 30, Application US/09733294A
4 Patent No. US2002045588A1
5 GENERAL INFORMATION:
6 APPLICANT: Brett P. Monia
7 APPLICANT: William Garde
8 APPLICANT: Susan M. Freier
9 APPLICANT: Edward V. Marcenicz
10 TITLE OF INVENTION: ANTISENSE MODULATION OF TERT EXPRESSION
11 FILE REFERENCE: 18PH-0527
12 CURRENT APPLICATION NUMBER: US/09/733,294A
13 CURRENT FILING DATE: 2000-12-07
14 PRIOR APPLICATION NUMBER: 09/572,423
15 PRIOR FILING DATE: 2000-05-16
16 NUMBER OF SEQ ID NOS: 108
17 SEQ ID NO 30
18 LENGTH: 51552
19 TYPE: DNA
20 ORGANISM: Homo sapiens
21 FEATURE:
22 NAME/KEY: exon
23 LOCATION: (1)...(11492)
24 OTHER INFORMATION: exon 1
25 NAME/KEY: intron
26 LOCATION: (11493)...(11596)
27 OTHER INFORMATION: intron 1
28 NAME/KEY: exon
29 LOCATION: (11597)...(12950)
30 OTHER INFORMATION: exon 2
31 NAME/KEY: intron
32 LOCATION: (12951)...(21566)
33 OTHER INFORMATION: intron 2
34 NAME/KEY: exon
35 LOCATION: (21567)...(21762)
36 OTHER INFORMATION: exon 3

```


1	NAME/KEY: int16on	LOCATION: (32851)
2	LOCATION: (32163)...	(32851)
3	OTHER INFORMATION: int16on	3
4	NAME/KEY: exon	
5	LOCATION: (32852)...	(24032)
6	OTHER INFORMATION: int16on	4
7	NAME/KEY: int16on	
8	LOCATION: (24033)...	(24719)
9	OTHER INFORMATION: int16on	4
10	NAME/KEY: exon	
11	LOCATION: (24720)...	(24699)
12	OTHER INFORMATION: int16on	5
13	NAME/KEY: int16on	
14	LOCATION: (24900)...	(25393)
15	OTHER INFORMATION: int16on	5
16	NAME/KEY: exon	
17	LOCATION: (25394)...	(25549)
18	OTHER INFORMATION: exon	6
19	NAME/KEY: int16on	
20	LOCATION: (25550)...	(30196)
21	OTHER INFORMATION: int16on	6
22	NAME/KEY: exon	
23	LOCATION: (30195)...	(30292)
24	OTHER INFORMATION: exon	7
25	NAME/KEY: int16on	
26	LOCATION: (30293)...	(31272)
27	OTHER INFORMATION: int16on	7
28	NAME/KEY: exon	
29	LOCATION: (31273)...	(31358)
30	OTHER INFORMATION: exon	8
31	NAME/KEY: int16on	
32	LOCATION: (31359)...	(33943)
33	OTHER INFORMATION: int16on	8
34	NAME/KEY: unsure	
35	LOCATION: 31450	
36	OTHER INFORMATION: unknown	
37	NAME/KEY: exon	
38	LOCATION: (33844)...	(33957)
39	OTHER INFORMATION: exon	9
40	NAME/KEY: int16on	
41	LOCATION: (33958)...	(35941)
42	OTHER INFORMATION: int16on	9
43	NAME/KEY: exon	
44	LOCATION: (35942)...	(36013)
45	OTHER INFORMATION: exon	10
46	NAME/KEY: int16on	
47	LOCATION: (36014)...	(37884)
48	OTHER INFORMATION: int16on	10
49	NAME/KEY: exon	
50	LOCATION: (37885)...	(38073)
51	OTHER INFORMATION: exon	11
52	NAME/KEY: int16on	
53	LOCATION: (38074)...	(41874)
54	OTHER INFORMATION: int16on	11
55	NAME/KEY: exon	
56	LOCATION: (41875)...	(42001)
57	OTHER INFORMATION: exon	12
58	NAME/KEY: int16on	
59	LOCATION: (42002)...	(42881)
60	OTHER INFORMATION: int16on	12
61	NAME/KEY: exon	
62	LOCATION: (42882)...	(42943)
63	OTHER INFORMATION: exon	13
64	NAME/KEY: int16on	
65	LOCATION: (42944)...	(46129)
66	OTHER INFORMATION: int16on	13
67	NAME/KEY: exon	
68	LOCATION: (46130)...	(46254)
69	OTHER INFORMATION: exon	14
70	NAME/KEY: int16on	
71	LOCATION: (46255)...	(47035)
72	OTHER INFORMATION: int16on	14
73	NAME/KEY: exon	

```

? LOCATION: (47036)..(47173)
? OTHER INFORMATION: exon 15
? NAME/KEY: intron
? LOCATION: (47174)..(47709)
? OTHER INFORMATION: intron 15
? NAME/KEY: exon
? LOCATION: (47710)..(50544)
? OTHER INFORMATION: exon 16
US-09-733-294A-30

```

Query Match 7.9%; Score 46.4; DB 3; Length 51552;
Best Local Similarity 50.4%; Pred. No. 0.0038;
Matches 139; Conservative 0; Mismatches 136; Indels 1; Gaps 1;

```

07      45  GAGGAGAGAGGTGGAGGTGAGTCCAGAGAGCCCTTCAAGATCAGACAGAGCCCTGCTCG 104
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 48820  GAGGCGAGGGGAGAGAGGGGGCCAGAGCCAGGCTTCCCATCTTCCCGCCACCCAGACAG 48765

```

Dc 48760 CCCCCCATCCCTGCAACCCAGTGGGGCCCCATCATCCTGCACCCTGCCA 48701

Accession	Sequence	Length
48700	GGCCCTTCATCATTCGCCGCGGCCAAGCTGGGCCCCCAGAGATCCCTGCTGCGCTTGGGCG	48641
224	CTTGCCTGCGCCCGCCGAGGCCCTCTCTTCTGGGCGAGAGATCTTCTCCCTGATGCGGACCA	283

48640 CTGGACTTACTGTTTATGTCTTCCAGGGATGGGGGCTCCCACTGTCTATCCCTACCTCCT 48581

Db 48580 TCCCTCTGCTCAGCATCAGAACTCCAGG 48545

US-10-184-644-202/c
; Sequence 202, Application US/10184644
; Publication No. US20030044930A1
; GENERAL INFORMATION:

```

: APPLICANT: Baker, Kevin P.
:
: APPLICANT: Chen, Jian
:
: APPLICANT: Desnoyers, Luc
:
: APPLICANT: Goddard, Audrey
:

```

```

? APPLICANT: Godowski, Paul U.
? APPLICANT: Gurney, Austin L.
? APPLICANT: Pan, James
? APPLICANT: Smith, Victoria
? APPLICANT: "Smith, William"

```

```

: APPLICANT: Wood, William I.
:
: APPLICANT: Zhang, Zemin
:
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
: TITLE OF INVENTION: ACIDS ENCODING THE SAME

```

; FILE REFERENCE: B3430RIC227
 ; CURRENT APPLICATION NUMBER: US/10/184,644
 ; CURRENT FILING DATE: 2002-06-28
 ; Prior Application removed - See File Wrapper or Palm

```

/ NUMBER OF SBR ID NOS: 012
/ SEC ID NO 202
/ LENGTH: 440
/ TYPE: PRT
/ ORIGINATOR: Home Garden

```

US-10-184-644-202	
Query Match	7.6%;
Best Local Similarity	Score 44.8; DB 5; Length 440;
	18.8%; Pred. No. 0.0058;

[illegible]

187 CAACAGCTGCAGGCTGCACCTTCGCGCTCCGCCCCCTGCCCTGCCCCCCCGAGAGCCCTTC 246

```
Db 329 CCCCCCDNCSGSSCASSGWS.SCSHCYCCSNSSCCSSGSSGSSGSSGSC 270
Qy 247 TTCTGGGCGAGAGATTCTCCTGTGACGACCATGCTATCTCTGAGGAGCTG 306
Db 269 SSSGNNHNCSSGSSGSS.SGSCSSGSSGSSGSSGSSGSSGSSGSSGSSGSC 210
Qy 307 GACACTTCATGATGATGAGAGCTGAGCCCTCCAGAAATCCAGTACTCCCTTGCTCCAG 366
Db 209 .ANAC.N..CCNCC.CW.TC..NKC.SCTSNK.RCDW.AC.CC.NC.C.CC.CC.SC.. 150
Qy 367 AATGAGTCCAGCCCGACCTGATCCAGTCTTCTAGAACTCTGAGCTCCCGTACTTG 426
Db 149 CDCGA.WTCSDC.BC.WSCYBTHTCY.BH.T.YC..DCANC.TDIT.CBYNC.THTTC 90
Qy 427 GGGGACTC 434
Db 89 .C.B.YBC 82
```

```
RESULT 11
US-10-184-634-202/c
; Sequence 202, Application US/10184634
; Publication No. US20030068684A1
GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P3430R1C217
CURRENT APPLICATION NUMBER: US/10/184,634
CURRENT FILING DATE: 2002-06-28
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 202
LENGTH: 440
TYPE: PRT
ORGANISM: Homo Sapien
US-10-184-634-202
```

```
Query Match 7.6%; Score 44.8; DB 5; Length 440;
Beat Local Similarity 18.8%; Pred. No. 0.0058;
Matches 58; Conservative 81; Mismatches 169; Indels 0; Gaps 0;
Qy 127 CGGAGCTGGGCCCCCGAGACCCGACCTCCGAGGAGCATGCTCATCAACACCTTC 186
Db 389 CSMSSC.CYRNHSCC..YNC.MS..YKNSBC.CY.C..CS.CSCT.NC.M.GC.MD 330
Qy 187 CAACAGCTGACGAGCTGACCTGCTGCTCCGCCCCCTGAGCCCGAGCCCTC 246
Db 329 CCCCCCDNCSGSSGSSGSS.SGSHCYCCSNSSGSSGSSGSSGSSGSSGSSGSC 270
Qy 247 TTCTGGGCGAGAGATTCTCCTGTGACGACCATGCTATCTCTGAGGAGCTG 306
Db 269 SSSGNNHNCSSGSSGSS.SGSCSSGSSGSSGSSGSSGSSGSSGSSGSSGSC 210
Qy 307 GACACTTCATGATGATGAGAGCTGAGCCCTCCAGAAATCCAGTACTCCCTTGCTCCAG 366
Db 209 .ANAC.N..CCNCC.CW.TC..NKC.SCTSNK.RCDW.AC.CC.NC.C.CC.CC.SC.. 150
Qy 367 AATGAGTCCAGCCCGACCTGATCCAGTCTTCTAGAACTCTGAGCTCCCGTACTTG 426
Db 149 CDCGA.WTCSDC.BC.WSCYBTHTCY.BH.T.YC..DCANC.TDIT.CBYNC.THTTC 90
```

```
Qy 427 GGGGACTC 434
Db 89 .C.B.YBC 82
```

```
RESULT 12
US-10-063-685-52/c
; Sequence 52, Application US/10063685
; Publication No. US20030180909A1
GENERAL INFORMATION:
; APPLICANT: Eaton, Dan L.
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P3230R1C1
CURRENT APPLICATION NUMBER: US/10/063,685
CURRENT FILING DATE: 2002-05-08
Prior Application removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 170
SEQ ID NO 52
LENGTH: 440
TYPE: PRT
ORGANISM: Homo Sapien
US-10-063-685-52
```

```
Query Match 7.6%; Score 44.8; DB 6; Length 440;
Beat Local Similarity 18.8%; Pred. No. 0.0058;
Matches 58; Conservative 81; Mismatches 169; Indels 0; Gaps 0;
Qy 127 CGGAGCTGGGCCCCCGAGACCCGACCTCCGAGGAGCATGCTCATCAACACCTTC 186
Db 389 CSMSSC.CYRNHSCC..YNC.MS..YKNSBC.CY.C..CS.CSCT.NC.M.GC.MD 330
Qy 187 CAACAGCTGACGAGCTGACCTGCTGCTCCGCCCCCTGAGCCCGAGCCCTC 246
Db 329 CCCCCCDNCSGSSGSSGSS.SGSHCYCCSNSSGSSGSSGSSGSSGSSGSSGSC 270
Qy 247 TTCTGGGCGAGAGATTCTCCTGTGACGACCATGCTATCTCTGAGGAGCTG 306
Db 269 SSSGNNHNCSSGSSGSS.SGSCSSGSSGSSGSSGSSGSSGSSGSSGSSGSC 210
Qy 307 GACACTTCATGATGATGAGAGCTGAGCCCTCCAGAAATCCAGTACTCCCTTGCTCCAG 366
Db 209 .ANAC.N..CCNCC.CW.TC..NKC.SCTSNK.RCDW.AC.CC.NC.C.CC.CC.SC.. 150
Qy 367 AATGAGTCCAGCCCGACCTGATCCAGTCTTCTAGAACTCTGAGCTCCCGTACTTG 426
Db 149 CDCGA.WTCSDC.BC.WSCYBTHTCY.BH.T.YC..DCANC.TDIT.CBYNC.THTTC 90
Qy 427 GGGGACTC 434
Db 89 .C.B.YBC 82
```

```
RESULT 13
US-10-123-155-10/c
; Sequence 10, Application US/10123155
; Publication No. US20030068794A1
GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Geo, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
```

```

; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C30
; CURRENT APPLICATION NUMBER: US/10/123,155
; PRIOR FILING DATE: 2002-04-15
; PRIOR APPLICATION removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 10
; LENGTH: 594
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-123-155-10

Query Match
Best Local Similarity 8.7%; Pred. No. 0.0061;
Matches 28; Conservative 135; Mismatches 160; Indels 0; Gaps 0;

7.6%; Score 44.8; DB 5; Length 594;
Query Match
Best Local Similarity 8.7%; Pred. No. 0.0061;
Matches 28; Conservative 135; Mismatches 160; Indels 0; Gaps 0;

QY 79 CAGAGCTACGACGACCCCTGCTCCGATCTCCCTAGACAAAGTCAGGAGCCTGGGC 138
DB 556 MM.M.CBT.STT.VA.M.YT.S.S.S.SYSYSYS.S.S.SDSYSYA.SYSYS.S.SWS 497
QY 139 CCCGAGACCCGAGCCTCCGAGCATGCTTCATCCATCAACACCTCCCAAGCTGCAG 198
DB 496 SYSYSSDDY.CYCCYRHHSDSYSYTY.CRCCTYT.SYRDCYHSCCSYCYTS 437
QY 199 GCTGCACTTCGCTGCTGCTCCGAGCATGCTTCATCCATCAACACCTCCCTGAGGAG 258
DB 436 SYSYRYSYSYSYSYSYSYTDYCSYRCCYYSYSSYSYSSAISTSSSSSS 377
QY 259 GAGGATTTCTCCCTGTCAGCACCAATGGCTCTATCTCAGAGAGCTGACACCTCCATG 318
DB 376 YYTSTNYC.T.CC...T.MCABCSYTTTTTTTT.HSCC.SA..A.M..YC.A.S 317
QY 319 GATGGAGCTGAGCCCTCCAGATCCAGATGATCCCTGCTGCTCCAGATCAAGTCCCA 378
DB 316 YSYSYS.SSS.S.SYMR.HRA.SHYTRS..S.MCY.YM.Y.YY.YSYSYCSHKT.. 257
QY 379 CCCGAGCTGATCCAGTCTTCTT 401
DB 256 ..TMTDM.T.T..MMY.KYB.H 234

RESULT 14
US-10-146-731-10/C
; Sequence 10, Application US/10146731
; Publication No. US20030129692A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
```

```

; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C323
; CURRENT APPLICATION NUMBER: US/10/146,731
; PRIOR FILING DATE: 2002-05-15
; PRIOR APPLICATION removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 10
; LENGTH: 594
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-146-731-10

Query Match
Best Local Similarity 8.7%; Pred. No. 0.0061;
Matches 28; Conservative 135; Mismatches 160; Indels 0; Gaps 0;

7.6%; Score 44.8; DB 6; Length 594;
Query Match
Best Local Similarity 8.7%; Pred. No. 0.0061;
Matches 28; Conservative 135; Mismatches 160; Indels 0; Gaps 0;

QY 79 CAGAGCTACGACGACCCCTGCTCCGATCTCCCTAGACAAAGTCAGGAGCCTGGGC 138
DB 556 MM.M.CBT.STT.VA.M.YT.S.S.S.SYSYSYS.S.S.SDSYSYA.SYSYS.S.SWS 497
QY 139 CCCGAGACCCGAGCCTCCGAGCATGCTTCATCCATCAACACCTCCCAAGCTGCAG 198
DB 496 SYSYSSDDY.CYCCYRHHSDSYSYTY.CRCCTYT.SYRDCYHSCCSYCYTS 437
QY 199 GCTGCACTTCGCTGCTGCTCCGAGCATGCTTCATCCATCAACACCTCCCTGAGGAG 258
DB 436 SYSYRYSYSYSYSYSYSYTDYCSYRCCYYSYSSYSYSSAISTSSSSSS 377
QY 259 GAGGATTTCTCCCTGTCAGCACCAATGGCTCTATCTCAGAGAGCTGACACCTCCATG 318
DB 376 YYTSTNYC.T.CC...T.MCABCSYTTTTTTTT.HSCC.SA..A.M..YC.A.S 317
QY 319 GATGGAGCTGAGCCCTCCAGATCCAGATGATCCCTGCTGCTCCAGATCAAGTCCCA 378
DB 316 YSYSYS.SSS.S.SYMR.HRA.SHYTRS..S.MCY.YM.Y.YY.YSYSYCSHKT.. 257
QY 379 CCCGAGCTGATCCAGTCTTCTT 401
DB 256 ..TMTDM.T.T..MMY.KYB.H 234

RESULT 15
US-10-140-472-10/C
; Sequence 10, Application US/10140472
; Publication No. US2003013888A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C168
; CURRENT APPLICATION NUMBER: US/10/140,472
; PRIOR FILING DATE: 2002-05-06
; PRIOR APPLICATION removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 10
```

; LENGTH: 594
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-140-472-10

Query Match 7.6%; Score 44.8; DB 6; Length 594;

Best Local Similarity 8.7%; Pred. No. 0.0061;
Matches 28; Conservative 135; Mismatches 160; Indels 0; Gaps 0;

```
QY CAGAGCTACCAAGACCCCTGCTCCGATCTCTAGACAAAGTCCAGGCGACGCTGGGC 138
   ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db MM.M.CBT.STT.YA.M.YT.S.S.S.SYSYSYS.S.S.SDSYSYA.SYSYS.S.SWS 497
   ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY CCCCAGACACCCAGCCTCCGAGGATGCTTCATTCATACACCCCTCCACAGCTGCAG 198
   ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db SYSYSSDDY.CYCCYRHHGSDSYSYY.CRCYYT.SYSRYDCHYSCCSDYYCY 437
   ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY GCTGCACTTCGCTGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCG 258
   ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db SYSRYYSYSYSYSYSYTDYCSYRCCYSSYSYSYSYSYSYSYSYSYSYSYS 377
   ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY GAGATTTCTCCCTGTCAGCCACATGCTCTATCTCCAGGAGCTGACACCTCCATG 318
   ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db YYYTSTNYC.T.CC...T.MCAABCSITTTTTTTT.HSCC.SA.A.M.YC.A.S 317
   ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY GATGGAGTGAAGCCCCCTCAGATCCAGTCACTCCCTTGAGCTCCAGATGAAGTCCA 378
   ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db YSYSYS.SSS.S.SYWR.HRA.SHYTBS.S.MYCY.YM.Y.Y.YSYSYCSRKM.. 257
   ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY CCCCAGCCTGATCCAGTCTTCTT 401
   ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db ..TMTDM.T.T..MNTY.KYB.H 234
   ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
```

Search completed: February 6, 2006, 11:43:45
Job time : 747 secs

November 2005

Published_Applications Nucleic Acid and Published_Applications Amino Acid database searches now generate two sets of results each. The Published_Applications databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.

Newly published applications will appear in the Published_Applications_New databases; older published applications make up the Published_Applications_Main databases.

- Searches run against Nucleic Acid Published_Applications produce two sets of results, with the extensions **.rnpbm** (Published_Applications_NA_Main) and **.rnpbn** (Published_Applications_NA_New).
- Searches run against Amino Acid Published_Applications produce two sets of results, with the extensions **.rapbm** (Published_Applications_AA_Main) and **.rapbn** (Published_Applications_AA_New).

This Page Blank (uspto)

Publication No. US2005025529A1
GENERAL INFORMATION:
APPLICANT: Bayer Healthcare AG
APPLICANT: Gold, Stefan
APPLICANT: Bruggemeier, Ulf
APPLICANT: Geerts, Andreas
TITLE OF INVENTION: Diagnostics and Therapeutics for Diseases Associated with a New
TITLE OF INVENTION: Human 5HT6 Receptor
FILE REFERENCE: laa 35 827
CURRENT APPLICATION NUMBER: US/10/502,893
CURRENT FILING DATE: 2004-07-27
PRIOR APPLICATION NUMBER: PCT/EP03/000479
PRIOR FILING DATE: 2003-01-20
PRIOR APPLICATION NUMBER: BP 02001942.8
PRIOR FILING DATE: 2002-02-01
NUMBER OF SEQ ID NOS: 5
SOFTWARE: PatentIn version 3.2
SEQ ID NO 1
LENGTH: 1984
TYPE: DNA
ORGANISM: Homo sapiens
US-11-502-893-1

Query Match	6.6%	Score 39	DB 7	Length 1984
Best Local Similarity	51.4%	Pred. No. 0.82		
Matches 90	Conservative 0	Mismatches 85	Indels 0	Gaps 0

QY	96	CGTGTTCGGACATCTCCCTGACGAAAGTCCAGGCGACGCTGGGCGCCCGAGACACCGCCT	155
Db	764	CCTGCTCTGGACCGCGCTTTCGACGTGATGTGCTGACGACGCGCTTCATCTTCAACTCTTGCGCT	823
QY	156	CGGACGGAATGTCCTCATCATTAACACCCCTCCAAAGCTGCAGGCTGCATCTTGCGCTGGC	215
Db	824	CATCAGCCTGAGACCGCTACCTGCTATTCCTCTTGCGCGCTGCGCTTAACAAGCTGCGCATGAC	883
QY	216	TCCGACCCCTGACGCTGACCCCGAGGCGCCCTCTTCTCTGCGGAGAGAGATTTCTCC	270
Db	884	GCCCTTGCGTGCCTTGAGCCCTTAATGCTCTGGAGCGGCTTGAGACGCTTGCGGCTTCCGC	938

```

RESULT 3
US-10-995-561-13274/c
Sequence 13274, Application US/10995561
Publication No. US20050272054A1
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele et al.
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
TITLE OF INVENTION: DETECTION AND USES THEREOF
FILE REFERENCE: CL001559
CURRENT APPLICATION NUMBER: US/10/995,561
CURRENT FILING DATE: 2004-11-24
NUMBER OF SEQ ID NOS: 85702
SOFTWARE: FastSeq For Windows Version 4.0
SEQ ID NO 13274
LENGTH: 415117
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(415117)
OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1
US-10-995-561-13274

```

Query Match	6.5%;	Score 38.2;	DB 7;	Length 415117;
Best Local Similarity	55.7%;	Pred. No. 9.6;		
Matches	73;	Conservative	0;	Mismatches 58;
				Indels 0;
				Gaps 0;

Oy 125 AGCGAGCCTGGGCCCCCGAAGACCCAGCTCTCGAGAGGATATTCTCATTCATAACACC 184
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3143 AGCAGAGCCGCGCGCAAGCCGCCAACCCGTTCCCCCACTTCCAAGCGCCAGAAGCGGCCCCC 308
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 185 TCCAACAGCTGCAGAGCTGACTTTCCGCTTGACTTCGCCGCCCTGTGCCCCCGAGACCCC 244

Db 3083 GCCCCCGGAGGCGCTCCCGCCCTGACACCCCTGCGCTTCGCGCCGCCCGC 3024
Qy 245 TCTTCTGGGC 255
Db 3023 GCGGCCCTGCC 3013

RESULT 4
US-11-000-688-628

; Sequence 628, Application US/110006888
; Publication No. US20050287544A1

```

? APPLICANT: BERTUCCI, Francots
? APPLICANT: HOUTGATTE, Remi
? APPLICANT: BURNEADW, Daniel
? TITLE OF INVENTION: GENE EXPRESSION PROFILING OF COLON CANCER WITH DNA ARRAYS
? FILE REFERENCE: 1423-R-03
? CURRENT APPLICATION NUMBER: US/11/000,688
? CURRENT FILING DATE: 2004-12-01
? PRIOR APPLICATION NUMBER: US 60/525,987
? PRIOR FILING DATE: 2003-12-01
? NUMBER OF SEQ ID NOS: 1596
? SOFTWARE: PatentIn version 3.2
? SEQ ID NO 628
? LENGTH: 2910
? TYPE: DNA
? ORGANISM: Artificial Sequence
? FEATURES:
? OTHER INFORMATION: Description of Artificial sequences:primer
? FEATURES:
? NAME/KEY: misc feature
? LOCATION: (1)..(2910)
? OTHER INFORMATION: cullin 2 (CUL2) gene.
? US-11-000-688-628

```

Query Match	6.4%	Score 37.6;	DB 8;	Length 2910;
Best Local Similarity	54.3%	Pred. No. 2.2;		
Matches 76;	Conservative 0;	Mismatches 64;	Indels 0;	Gaps 0;

QY 130 AGCCCTGGGCCCCCAGAGCACCCAGCCTCCGAGGACATGTCCTATGCATTAACAACCCCTCGAA 189
 Db 54 AGCAGGGGAAACCGCGACTGACAGCCGCGCGCCCGCCGCGCTTCGAGCCACCTTCTTGACCG 113
 QY 190 CAGCTGCAGGCTGCATTCGCTGCTGCCGCGCCCTGCCTGCGCCCCCGAGGCCCTCTTC 249
 Db 114 GGGCTTCGTCTTACTACTCTTCGGGCTGCTCCCTCTGTCCTCCCTGCTGCCCTTTCGCC 173
 QY 250 CTGGGCGAGGAGATTTCTC 269
 Db 174 CTGCTTCTGCAGAAATTC 193

RESULT

US-10-995-561-66759

```
; Sequence 66759, Application US/10995561
; Publication No. US20050272054A1
```

```

/ APPLICANT: CARGILL, Michele et al.
/ TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
/ TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
/ TITLE OF INVENTION: DETECTION AND USES THEREOF
/ FILE REFERENCE: C1001559
/ CURRENT APPLICATION NUMBER: US/10/995,561
/ CURRENT FILING DATE: 2004-11-24
/ NUMBER OF SEQ ID NOS: 85702
/ SOFTWARE: FASTSEQ for Windows Version 4.0
/ SEQ ID NO 66759
/ LENGTH: 201
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-995-561-66759

```



```
Query Match      6.3%; Score 37.4; DB 7; Length 201;
Best Local Similarity 54.0%; Pred. No. 0.93;
Matches 74; Conservative 1; Mismatches 62; Indels 0; Gaps 0;

QY 138 CCCCCGAGACCCAGCTCCGCGAGCATGTCATCATCAATAACCTCCAAAGAGTGCA 197
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3 CCGGCCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 62

QY 198 GGGTCGACTTCGCTGAGTCCGCCCTGCGCCCGCGAGCCCTCTTCTTGAGGCGCA 257
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 63 GCGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 122

QY 258 GAGAGATTCTCCCTGT 274
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 123 GCGCGAGCTCTGCGCGT 139

RESULT 6
US-10-995-561-13385
; Sequence 13385, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; PRIOR FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for windows Version 4.0
; SEQ ID NO 13385
; LENGTH: 18138
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-13385

Query Match      6.3%; Score 37.4; DB 7; Length 18138;
Best Local Similarity 54.0%; Pred. No. 4.9;
Matches 74; Conservative 1; Mismatches 62; Indels 0; Gaps 0;

QY 138 CCCCCGAGACCCAGCTCCGCGAGCATGTCATCATCAATAACCTCCAAAGAGTGCA 197
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 6274 CCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 6333

QY 198 GGGTCGACTTCGCTGAGTCCGCCCTGCGCCCGCGAGCCCTCTTCTTGAGGCGCA 257
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 6334 GCGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 6393

QY 258 GAGAGATTCTCCCTGT 274
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 6394 GCGCGAGCTCTGCGCGT 6410

RESULT 7
US-10-821-234-596
; Sequence 596, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; PRIOR FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pc_seq_genes Version 1.0
; SEQ ID NO 596
; LENGTH: 2139
```

```
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-821-234-596

Query Match      6.3%; Score 37.2; DB 7; Length 2139;
Best Local Similarity 59.4%; Pred. No. 2.5;
Matches 63; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 218 CCGCCCCCTGCCCCCGCGAGCCCTCTTCTTGAGGCGAGAGATTCTCCCTGTGAG 277
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 408 CCAAGCGGCTCTCTTCTCCGCGCCACTTTCAGAGGCCAGAGAGCTCCCGCGAGCCAG 467

QY 278 CCACCATTTGCTTATCTCAGAGAGCTGAGACCTTCATGATGG 323
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 468 CCCATGCTGCTGAGGGGTCCAGAGGGGCTGAGAGCTGCTGCTGATG 513

RESULT 8
US-11-136-527-3584
; Sequence 3584, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; PRIOR FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 3584
; LENGTH: 2792
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-3584

Query Match      6.3%; Score 37; DB 8; Length 2792;
Best Local Similarity 52.2%; Pred. No. 3.2;
Matches 82; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 347 TGAATCCCTTGGGCTCCAGAAATGAGTCCACCCAGGCTGATTCAGTCTTGTAGAG 406
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1134 TGCACCGGCTGTCTCTCCGGGCTCCCGAGCGCCCTGCGCGGTCTCAGTCTCGCTGAG 1193

QY 407 CTCTGAGCTCCGGTACTTGGGGGACTTGGGCTGATGATCTTTCTTGACATTGACA 466
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1194 GTCCCTCGCCAGTCTTGTGATGATCCACCAGAGAGGCTCTGCTGCTGCTGGAAG 1253

QY 467 CATCTGCGGTAGAAAGAGGCTGACGCGGCCCCACC 503
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1254 CAGCGAAGCGAGAGGGGCGCACGCGAGAGCCCC 1290

RESULT 9
US-11-136-527-150/C
; Sequence 150, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; PRIOR FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 150
; LENGTH: 3958
; TYPE: DNA
```

ORGANISM: Rattus norvegicus
US-11-136-527-150

Query Match 6.3%; Score 37; DB 8; Length 3958;
Best Local Similarity 53.0%; Pred. No. 3.6;
Matches 79; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 79 CAGAGCTACGACGAGCCCTGCTCCGATCTCCCTAGACAAAGTCACAGCGAGCTGGGC 138
DB 2072 CAGAACTTCACTTGCCCTTGCCCTGCTCATCTTGAGCATCTTGTCTCATAGCTGATGTC 2013
QY 139 CCCCAGACACCGAGCTCCGAGGAGATGCTCATTCATACACCTTCCAGAGCTGCAG 198
DB 2012 AAAGCCCCAGAGAGCTTGACGAGGAGCTGCTGCTCTTCTCTCTCCAGGTTCTT 1953
QY 199 GGTGCACTTCGCTGCTCCGAGCCCTGC 227
DB 1952 GAGGCACTTCTGTCAGCGGAGCACTTGC 1924

RESULT 10

US-11-121-086-5/C
Sequence 5, Application US/11121086
Publication No. US20050266459A1
GENERAL INFORMATION:
APPLICANT: NIELSEN, KIRSTEN V.
APPLICANT: POULSEN, TIM S.
TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
FILE REFERENCE: 09138,6000-00000
CURRENT FILING DATE: 2005-05-04
PRIOR APPLICATION NUMBER: 60/567,570
NUMBER OF SEQ ID NOS: 107
SOFTWARE: PatentIn version 3.3
SEQ ID NO 5
LENGTH: 153376
TYPE: DNA
ORGANISM: Homo sapiens
US-11-121-086-5

Query Match 6.2%; Score 36.8; DB 8; Length 153376;
Best Local Similarity 51.9%; Pred. No. 16;
Matches 83; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 96 CCTGCTCCGATCTCCCTAGACAAAGTCACAGCGAGCTGGGCCCCGAGACCCAGCCT 155
DB 16325 CCT 16286
QY 156 CCGCAGGAGATGCTCATTCATACACCTTCCAGAGCTGAGGCTGACCTTGGCTGGC 215
DB 16285 CCGTCT 16206
QY 216 TCCGCCCCCTGCTGAGCCCCCGAGGCCCTTCTTCTGGGC 255
DB 16205 CCT 16166

RESULT 11

US-10-858-730-170
Sequence 170, Application US/10858730
Publication No. US20050255568A1
GENERAL INFORMATION:
APPLICANT: Bailey, Richard B.
APPLICANT: Blomquist, Paul
APPLICANT: Doten, Reed
APPLICANT: Driggers, Edward M.
APPLICANT: Madden, Kevin T.
APPLICANT: O'Leary, Jessica
APPLICANT: O'Toole, George
APPLICANT: Trueheart, Joshua
APPLICANT: Walbridge, Michael J.
APPLICANT: Yorgey, Peter S.

TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR AMINO ACID
FILE OF INVENTION: PRODUCTION
FILE REFERENCE: 14184-030001
CURRENT APPLICATION NUMBER: US/10/858,730
CURRENT FILING DATE: 2004-06-01
PRIOR APPLICATION NUMBER: US 60/475,000
PRIOR FILING DATE: 2003-05-30
PRIOR APPLICATION NUMBER: US 60/551,860
PRIOR FILING DATE: 2004-03-10
NUMBER OF SEQ ID NOS: 364
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 170
LENGTH: 1161
TYPE: DNA
ORGANISM: Streptomyces coelicolor
US-10-858-730-170

Query Match 6.2%; Score 36.4; DB 7; Length 1161;
Best Local Similarity 48.1%; Pred. No. 3.3; 111; Indels 0; Gaps 0;
Matches 103; Conservative 0; Mismatches 111; Indels 0; Gaps 0;

QY 96 CCTGCTCCGATCTCCCTAGACAAAGTCACAGCGAGCTGGGCCCCGAGACCCAGCCT 155
DB 468 CCGCTGCTCCGATCAGACGATGCGGAGGTGCCAGAGTCCCGGAGCGCGGC 527
QY 156 CCGCAGGAGATGCTCATTCATACACCTTCCAGAGCTGAGGCTGCACTTGGCTGGC 215
DB 528 CCGGCTCGTCTGACACACCTTGCACCCCGTACCTCCAGAGCGCGCTGGCCCTGG 587
QY 216 TCCGCCCCCTGCTGAGCCCCCGAGGCCCTTCTTCTGGGGCGAGAGATTCTCCCTGC 275
DB 588 CGCGAGCTGCTGCGACACTGCTGACCAAGTACAGAGGAGGAGGAGCTGAGG 647
QY 276 AGCCACATTGGCTCTATCTCAGGAGCTGAGC 309
DB 648 CGCGCGCTGATCTGAGGCGAGACGAGACTGGC 681

RESULT 12

US-10-714-887-113
Sequence 113, Application US/10714887
Publication No. US20060015972A1
GENERAL INFORMATION:
APPLICANT: Mendel Biotechnology, Inc.
APPLICANT: HEARD, Jacqueline
APPLICANT: RICHMANN, Jose Luis
APPLICANT: CREELMAN, Robert
APPLICANT: RATCLIFFE, Oliver
APPLICANT: CANALES, Roger
APPLICANT: REPETTI, Peter
APPLICANT: KUMIMOTO, Roderick W
APPLICANT: GUTTERSON, Neal
APPLICANT: REUBER, T. Lynne
APPLICANT: PINEDA, Omaira
APPLICANT: SHERMAN, Bradley K
TITLE OF INVENTION: PLANT TRANSCRIPTIONAL REGULATORS OF DROUGHT STRESS
FILE REFERENCE: MB10058-CIP
CURRENT APPLICATION NUMBER: US/10/714,887
CURRENT FILING DATE: 2003-11-13
PRIOR APPLICATION NUMBER: 10/412,699
PRIOR FILING DATE: 2003-04-10
PRIOR APPLICATION NUMBER: 09/506,720
PRIOR FILING DATE: 2000-02-17
PRIOR APPLICATION NUMBER: 60/135,134
PRIOR FILING DATE: 1999-05-20
PRIOR APPLICATION NUMBER: 09/394,519
PRIOR FILING DATE: 1999-09-13
PRIOR APPLICATION NUMBER: 09/533,392
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: 09/533,029
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: 09/532,591
PRIOR FILING DATE: 2000-03-22

PRIOR APPLICATION NUMBER: 09/533,030
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: 60/125,814
PRIOR FILING DATE: 1999-03-23
PRIOR APPLICATION NUMBER: 09/713,994
PRIOR FILING DATE: 2000-11-16
Remaining prior application data removed - See file Wrapper or PALM.
NUMBER OF SEQ ID NOS: 430
SOFTWARE: PatentIn version 3.2
SEQ ID NO: 113
LENGTH: 1529
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: G3680 Predicted polypeptide sequence is orthologous to G2999
US-10-714-887-113

Query Match 6.2%; Score 36.4; DB 6; Length 1529;
Best Local Similarity 48.5%; Pred. No. 3.6;
Matches 100; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

QY 330 GCCCCTCAGATCCAGATCTCCCTTGGCTCCAGATGAGGCCACCCAGCTCA 389
DB 405 GCCCGCGCGCGCGCTGGCGCTCCCGCTCCGCGAGCTGATGACAGGCCGCA 464
QY 390 TCCAGTCTTTAGAGCTCTGAGCTCCCGTACTTGGGGGAACTGGGCTGATGACTT 449
DB 465 CCGCGCGCGGAGAGAGCGCCGAGAGACCCGCGGGGCTGTGAGCGCCAGCATC 524
QY 450 CTTTCTGACATGACATCTGCGGTAGAAAAGAGCTGACGCGGCCCAACAGAGCC 509
DB 525 CGATTCCGACCTCGAAGGCTCGAGATGACAGAGAGCGGTGCTGCTCCCGCGCGCC 584
QY 510 TCCTCAACAACCTCTTCTGTGCCCCAG 535
DB 585 GCCGCAACCACTGCTGCGCGCGCCG 610

RESULT 13
US-11-121-086-101/c
Sequence 101, Application US/11121086
Publication No. US20050266459a1
GENERAL INFORMATION:
APPLICANT: POULSEN, TIM S.
TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
FILE REFERENCE: 09138, 6000-00000
CURRENT APPLICATION NUMBER: US/11/121,086
CURRENT FILING DATE: 2005-05-04
PRIOR APPLICATION NUMBER: 60/567,570
PRIOR FILING DATE: 2004-05-04
NUMBER OF SEQ ID NOS: 107
SOFTWARE: PatentIn version 3.3
SEQ ID NO: 101
LENGTH: 185393
TYPE: DNA
ORGANISM: Homo sapiens
US-11-121-086-101

Query Match 6.1%; Score 36; DB 8; Length 185393;
Best Local Similarity 51.9%; Pred. No. 27;
Matches 81; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 120 AGTCAGCGCAGCCTGGGCGCCCGAGACACCACTCCGAGGAGCATGCTTCATCCATTA 179
DB 143649 AGCGGGGCAAGCCGAGCCCGAGAGCGCGGAGCAGCAGCCCTCTCCAGC 143590
QY 180 CACCTCCAGCAGCTGCGAGCTGCACTTGCGGTGCTCCGCGCCCTGCGCCCGCGCA 239
DB 143589 GCGCTCCCTGCGCGCGCGCGCAAGCCCTCTGCGCTCCCTCCCGCGCGCGCG 143530
QY 240 GCCCTCTTCTCTGCGGAGAGGATTTCTCCCTGTC 275

DB 143529 GTCTCGCTTCTTCCCGCGCGCTGCGCCCTGCTC 143494

RESULT 14
US-11-153-238-6/c
Sequence 6, Application US/11153238
Publication No. US20060019281a1
GENERAL INFORMATION:
APPLICANT: KILLARY, ANN
APPLICANT: LOTT, STEVE
APPLICANT: CHANDLER, DAWN
TITLE OF INVENTION: THE TUMOR SUPPRESSOR CAR-1
FILE REFERENCE: UTSC:65105
CURRENT APPLICATION NUMBER: US/11/153,238
CURRENT FILING DATE: 2005-06-15
PRIOR APPLICATION NUMBER: US/09/927,091
PRIOR FILING DATE: 2001-08-09
PRIOR APPLICATION NUMBER: 60/227,560
PRIOR FILING DATE: 2000-08-23
PRIOR APPLICATION NUMBER: 60/225,033
PRIOR FILING DATE: 2000-08-10
NUMBER OF SEQ ID NOS: 9
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO: 6
LENGTH: 45845
TYPE: DNA
ORGANISM: Human
US-11-153-238-6

Query Match 6.1%; Score 35.8; DB 8; Length 45845;
Best Local Similarity 56.3%; Pred. No. 18;
Matches 67; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 142 CGAGACCCAGCCTCCGAGAGCATCTCTCATTCATTAACCCCTCAACAGCTGAGCT 201
DB 18203 CGAGACCCAGCCTCTGAGAGCTGCAATCTCTGCGCGCAGGAGCCCTTGAAGCT 18144
QY 202 GCATTCGCTGAGCTCCGCGCCCTGCGCCCGGAGCCCTCTTCTGCGGCGAGGA 260
DB 18143 GCATCTCTCTGCGGCGAGAGACCACTGACAGCTGCAATCTCTGCGGCGAGCA 18085

RESULT 15
US-11-043-752-3912
Sequence 3912, Application US/11043752
Publication No. US20060014165a1
GENERAL INFORMATION:
APPLICANT: Hakonarson, Hakon
APPLICANT: Gurney, Mark E.
APPLICANT: Haladi, Eva
TITLE OF INVENTION: METHODS OF DIAGNOSIS AND TREATMENT FOR
TITLE OF INVENTION: ASTHMA AND OTHER RESPIRATORY DISEASES BASED ON HAPLOTYPE
FILE REFERENCE: 2345, 2044-003
CURRENT APPLICATION NUMBER: US/11/043,752
CURRENT FILING DATE: 2005-01-26
PRIOR APPLICATION NUMBER: PCT/US04/022446
PRIOR FILING DATE: 2004-07-14
PRIOR APPLICATION NUMBER: 60/487,072
PRIOR FILING DATE: 2003-07-14
PRIOR APPLICATION NUMBER: 60/559,611
PRIOR FILING DATE: 2004-04-05
NUMBER OF SEQ ID NOS: 4326
SOFTWARE: fastSeq for Windows Version 4.0
SEQ ID NO: 3912
LENGTH: 244
TYPE: DNA
ORGANISM: Homo sapiens
US-11-043-752-3912

Query Match 6.0%; Score 35.4; DB 8; Length 244;
Best Local Similarity 51.6%; Pred. No. 3.4;
Matches 81; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 6, 2006, 09:53:16 ; Search time 200 Seconds

(without alignments)
430.591 Million cell updates/sec

Title: US-10-069-386a-2

Perfect score: 1033
Sequence: 1 MEGGLKRRKHSLEEEERNE.....APGSWEWELDHIMEILLGS 196

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_21.*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1033	100.0	196	AA35402	Aab35402 Replicat
2	1026	99.3	196	AA93922	Aam93922 Human pol
3	1026	99.3	196	ADJ32052	Adl32052 Human pro
4	1026	99.3	196	ADZ70576	Adz70576 Human pro
5	791	76.6	157	AA674742	Aag74742 Human col
6	645.5	62.5	142	ABO60075	Aboc60075 Human gen
7	402	38.9	111	ABO60074	Aboc60074 Human gen
8	216	20.9	236	AA36004	Aay36004 Extended
9	216	20.9	236	ADP19312	Adp19312 Human sec
10	213	20.6	236	AA44362	Aay44362 Human cel
11	213	20.6	236	AA933724	Aam933724 Human pol
12	213	20.6	236	AA689292	Aag89292 Human sec
13	213	20.6	236	ADL1644	Adl1644 Human pro
14	213	20.6	236	AA25550	Aam25550 Human pro
15	211	20.4	236	AA31829	Aay31829 Human adu
16	205	18.8	222	ABO60367	Aboc60367 Human gen
17	205	18.8	236	AA60367	Aay60367 Human gen
18	136	13.2	237	AAO22897	Aao22897 Mouse hae
19	125	12.1	314	ABP65076	Abp65076 Hypoxia-1
20	125	12.1	314	ADR14626	Adr14626 Human NF-
21	125	12.1	314	ADX06910	Adx06910 Cyclin-de
22	106	10.3	241	AA39725	Aam39725 Human pol
23	106	10.3	254	AA58258	Aab58258 Lung canc
24	106	10.3	254	AA673682	Aag73682 Human col

25	106	10.3	254	4	AA41511	Aam41511 Human pol
26	105.5	10.2	578	4	AA38707	Aam38707 Human pol
27	105.5	10.2	620	7	ADN95361	Adn95361 Human BBC
28	105.5	10.2	620	8	ADP54705	Adp54705 Human PRO
29	105.5	10.2	650	4	AA38706	Aam38706 Human PRO
30	105.5	10.2	685	7	ADZ828201	Adz828201 Human KDD
31	105.5	10.2	759	4	AA40492	Aam40492 Human pol
32	105.5	10.2	759	4	AA40493	Aam40493 Human pol
33	105.5	10.2	760	7	ADN95128	Adn95128 Human LBC
34	105	10.2	241	5	AB806375	Abb06375 Human CHD
35	105	10.2	241	5	AAO22898	Aao22898 Human hae
36	105	10.2	241	5	ABM81593	Abm81593 Tumour-as
37	105	10.2	241	9	ADY17410	Ady17410 PRO polyp
38	104.5	10.1	740	4	ABB11713	Abb11713 Human KIA
39	104.5	10.1	879	5	ABP69285	Abp69285 Human pro
40	104	10.1	322	4	ABG32868	Abg32868 Novel hum
41	101.5	9.8	728	7	ADD46841	Add46841 Human pro
42	101.5	9.8	728	7	AD56294	Ad56294 Human pro
43	101.5	9.8	728	8	ADJ66562	Adj66562 P13 Kinase
44	98	9.5	562	7	ADB65060	Adb65060 Human pro
45	98	9.5	562	8	ADR58973	Adr58973 Human Elk

ALIGNMENTS

RESULT 1
AAB35402
ID AAB35402 standard; protein; 196 AA.
XX
AC AAB35402;
XX
DT 23-MAY-2001 (first entry)
XX
DE Replication protein A binding transcriptional activator 1 RBT1.
XX
KW RBT1; replication protein A binding transcriptional activator 1; RPA32;
XX gene therapy; apoptosis; cancer; leukemia.
XX
OS Unidentified.
XX
PN WO200114546-A2.
XX
PD 01-MAR-2001.
XX
PF 17-AUG-2000; 2000WO-CA000948.
XX
PR 19-AUG-1999; 99US-0149472P.
XX
PA (TRAN-) CENT TRANSLATIONAL RES IN CANCER.
XX
PI Alaoui-Jamali MA, Cho JM;
XX
DR WPI; 2001-218447/22.
XX
DR N-PSDB; AAF28052.
XX
PT Novel replication protein A binding transcriptional activator 1 gene,
XX useful for treating neoplastic disorders such as cancer and in gene
XX therapy.
XX
PS Disclosure; Fig 1; 16pp; English.
XX
CC The present invention provides the protein and coding sequences of the
XX replication protein A binding transcriptional activator 1 (RBT1). The
XX protein is capable of inducing apoptosis. The sequences are useful in the
XX gene therapy and other methods of treatment of cancer, including
XX leukaemias. The present sequence is the RBT1 protein
XX
SQ Sequence 196 AA;
XX
Query Match 100.0%; Score 1033; DB 4; Length 196;
Best Local Similarity 100.0%; Pred. No. 1.9e-92;
Matches 196; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEGGKKRKHSDLEEEEREMWSPAGLQSYQOALLRISLDKYQRSIGPPAPSLRRHVLIN 60
DB 1 MEGGKKRKHSDLEEEEREMWSPAGLQSYQOALLRISLDKYQRSIGPPAPSLRRHVLIN 60
QY 61 TLQOLQALRLAPALPPEPLFLGEBDFSLSATIGSLIRELDTSMOETEPQNPVTPLG 120
DB 61 TLQOLQALRLAPALPPEPLFLGEBDFSLSATIGSLIRELDTSMOETEPQNPVTPLG 120
QY 121 LQNEVPQPDVPFLFALSSRYLGDGGLDDFLDIDTSAVEKEPARAPPEPPHNLFCAPGS 180
DB 121 LQNEVPQPDVPFLFALSSRYLGDGGLDDFLDIDTSAVEKEPARAPPEPPHNLFCAPGS 180
QY 181 WENNELDHIMEIILIGS 196
DB 181 WENNELDHIMEIILIGS 196

RESULT 2
AAM93922
ID AAM93922 standard; protein; 196 AA.

AC AAM93922;
DT 06-NOV-2001 (first entry)
DE Human polypeptide, SEQ ID NO: 4085.
KM Human, full length cDNA; cDNA synthesis; oligo-capping.
OS Homo sapiens.
PN EP130094-A2.
PD 05-SEP-2001.
PF 07-JUL-2000; 2000EP-00114089.
PR 08-JUL-1999; 99JP-00194486.
PR 11-JAN-2000; 2000JP-00118774.
PR 02-MAY-2000; 2000JP-00183765.
PA (HELI-) HELIX RES INST.
PI Ota T, Nishikawa T, Isegai T, Hayashi K, Ishii S, Kawai Y,
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H,
DR WPI; 2001-524255/58.
DR N-PSDB; AAK94884.
PT 830 Primers useful for synthesizing full length cDNA clones and their use
PT in genetic manipulation.
PS Claim 8; SEQ ID NO 4085; 1380pp + Sequence listing; English.
XX The invention relates to primers for synthesizing full length cDNA
XX clones. 830 cDNA molecules encoding a human protein have been isolated
XX and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
XX been determined. Primers for synthesizing the full length cDNA are useful
XX for clarifying the function of the protein encoded by the cDNA. The full
XX length clones were obtained by construction of full length enriched cDNA
XX libraries that were synthesized by the oligo-capping method. The primers
XX enable the production of the full length cDNA easily without any special
XX methods. The present sequence is a polypeptide encoded by a full length
XX human cDNA of the invention. Note: The sequence data for this patent did
XX not form part of the printed specification, but was obtained in CD-ROM
XX format directly from EPO
SQ Sequence 196 AA;

Query Match 99.3%; Score 1026; DB 4; Length 196;
Best Local Similarity 99.5%; Pred. No. 9e-92;
Matches 195; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MEGGKKRKHSDLEEEEREMWSPAGLQSYQOALLRISLDKYQRSIGPPAPSLRRHVLIN 60
DB 1 MEGGKKRKHSDLEEEEREMWSPAGLQSYQOALLRISLDKYQRSIGPPAPSLRRHVLIN 60
QY 61 TLQOLQALRLAPALPPEPLFLGEBDFSLSATIGSLIRELDTSMOETEPQNPVTPLG 120
DB 61 TLQOLQALRLAPALPPEPLFLGEBDFSLSATIGSLIRELDTSMOETEPQNPVTPLG 120
QY 121 LQNEVPQPDVPFLFALSSRYLGDGGLDDFLDIDTSAVEKEPARAPPEPPHNLFCAPGS 180
DB 121 LQNEVPQPDVPFLFALSSRYLGDGGLDDFLDIDTSAVEKEPARAPPEPPHNLFCAPGS 180
QY 181 WENNELDHIMEIILIGS 196
DB 181 WENNELDHIMEIILIGS 196

RESULT 3
ADL32052
ID ADL32052 standard; protein; 196 AA.

AC ADL32052;
DT 20-MAY-2004 (first entry)
DE Human protein encoded by a full length cDNA clone SeqID 4085.
KM human; medicine; signal transduction; glycoprotein; transcription;
KM oligo-capping method.
OS Homo sapiens.
PN EP1396543-A2.
PD 10-MAR-2004.
PF 07-JUL-2000; 2003EP-00025638.
PR 08-JUL-1999; 99JP-00194486.
PR 11-JAN-2000; 2000JP-00118774.
PR 02-MAY-2000; 2000JP-00183865.
PR 07-JUL-2000; 2000EP-00114089.
PA (REAS-) RES ASSOC BIOTECHNOLOGY.
PI Ota T, Nishikawa T, Isegai T, Hayashi K, Ishii S, Kawai Y,
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H,
DR WPI; 2004-204755/20.
DR N-PSDB; ADL32051.
PT New oligonucleotide primers (830 cDNAs) useful for synthesizing full
PT length human cDNAs.
PS Example 1; SEQ ID NO 4085; 1340pp; English.
XX This invention relates to a novel primers useful for synthesizing full
XX length cDNA molecules that encode human proteins. Specifically, it refers
XX to secretory or membrane proteins that are potential therapeutic agents/
XX target molecules in the field of medicine, and in particular genes
XX encoding proteins that are associated with signal transduction,
XX glycoproteins and transcription. The present invention describes a method
XX for efficiently cloning a full length human cDNA from both the 5' and 3'
XX ends using the oligo-capping method. This polypeptide sequence is a full
XX length human protein of the invention.
SQ Sequence 196 AA;

Query Match 99.3%; Score 1026; DB 8; Length 196;
Best Local Similarity 99.5%; Pred. No. 9e-92;
Matches 195; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 MEGGKRRKHSDEEEERWEMSPAGLQSYQALLRLSLDKVORSIGPPAPSLRRHYLIHN 60
DB 1 MWGGLKRRKHSDEEEERWEMSPAGLQSYQALLRLSLDKVORSIGPPAPSLRRHYLIHN 60
QY 61 TLQQLQALLRLAPALPPEPLFLGEBDFSLSATIGSIILRELDTSMGTEPPQNPTPIG 120
DB 61 TLQQLQALLRLAPALPPEPLFLGEBDFSLSATIGSIILRELDTSMGTEPPQNPTPIG 120
QY 121 LQNEVPQPDVPVFLBALSSRYLGDSDGLDDFLDITDTSAVEKEPARAPPEPPHNLFCAPGS 180
DB 121 LQNEVPQPDVPVFLBALSSRYLGDSDGLDDFLDITDTSAVEKEPARAPPEPPHNLFCAPGS 180
QY 181 MWMNELDHIMEIILIGS 196
DB 181 MWMNELDHIMEIILIGS 196

RESULT 4
ADZ70576 standard; protein; 196 AA.
AC ADZ70576;
DT 30-JUN-2005 (first entry)
DE Human protein from lung cancer marker gene RBT1.
KW Tumor marker; lung tumor; cytostatic; neoplasm; expression;
KN DNA microarray.
XX Homo sapiens.
OS WO2005032495-A2.
XX 14-APR-2005.
XX 01-OCT-2004; 2004WO-US034163.
XX 03-OCT-2003; 2003US-0508355P.
XX (FARB ) BAYER PHARM CORP.
XX Taylor I, Pauloeki NR, Bigwood D;
XX MPI; 2005-285325/29.
XX N-PSDB; ADZ70575.
XX Providing a patient diagnosis for lung cancer comprises comparing the
XX level of expression of genes or gene products in a biological sample from
XX the patient with that from a normal individual.
XX Claim 3; SEQ ID NO 261; 60pp; English.
XX
XX The invention relates to providing a patient diagnosis for lung cancer
XX comprising comparing the level of expression of genes or gene products in
XX a biological sample from the patient with the level of expression of
XX genes or gene products in a biological sample from a normal individual.
XX Also included are distinguishing between normal and disease tissues,
XX monitoring the response of a patient being treated for lung cancer by
XX administering an anti-cancer agent, identifying a compound useful for the
XX treatment of lung cancer and an array for distinguishing between normal
XX and disease tissues (comprising 2 or more probes corresponding to 2 or
XX more genes selected from any of the 200 nucleotide sequences given in the
XX specification, or 2 or more polypeptides comprising any of the 200 amino
XX acid sequences given in the specification). In providing a patient
XX diagnosis for lung cancer, one or more genes are selected from any of the
XX 200 nucleotide sequences as mentioned in the specification, or one or
XX more gene products are polypeptides selected from any of the 20 amino
XX acid sequences mentioned in the specification. The methods are useful for
XX detecting and treating lung cancer. These may also be used for designing,
XX identifying and optimizing therapeutics for cancer. The present sequence
XX represents a protein from one of the 200 lung cancer marker genes. Note:
XX The sequence data for this patent did not form part of the printed
```

```
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 196 AA;
Query Match 99.3%; Score 1026; DB 9; Length 196;
Best Local Similarity 99.5%; Pred. No. 96-92;
Matches 195; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MEGGKRRKHSDEEEERWEMSPAGLQSYQALLRLSLDKVORSIGPPAPSLRRHYLIHN 60
DB 1 MWGGLKRRKHSDEEEERWEMSPAGLQSYQALLRLSLDKVORSIGPPAPSLRRHYLIHN 60
QY 61 TLQQLQALLRLAPALPPEPLFLGEBDFSLSATIGSIILRELDTSMGTEPPQNPTPIG 120
DB 61 TLQQLQALLRLAPALPPEPLFLGEBDFSLSATIGSIILRELDTSMGTEPPQNPTPIG 120
QY 121 LQNEVPQPDVPVFLBALSSRYLGDSDGLDDFLDITDTSAVEKEPARAPPEPPHNLFCAPGS 180
DB 121 LQNEVPQPDVPVFLBALSSRYLGDSDGLDDFLDITDTSAVEKEPARAPPEPPHNLFCAPGS 180
QY 181 MWMNELDHIMEIILIGS 196
DB 181 MWMNELDHIMEIILIGS 196

RESULT 5
AAG74742 standard; protein; 157 AA.
ID AAG74742;
AC AAG74742;
DT 03-SEP-2001 (first entry)
DE Human colon cancer antigen protein SEQ ID NO:5506.
KW Human; colon cancer; colon cancer antigen; diagnosis; detection;
KN colorectal carcinoma.
XX Homo sapiens.
OS WO200122920-A2.
XX 05-APR-2001.
XX 28-SEP-2000; 2000WO-US026524.
XX 29-SEP-1999; 99US-0157137P.
XX 03-NOV-1999; 99US-0163280P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Ruben SM, Barash SC, Birse CE, Rosen CA;
XX MPI; 2001-23537/24.
XX N-PSDB; AAH34147.
XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
XX useful for preventing, diagnosing and/or treating colorectal cancers.
XX Claim 11; Page 7098-7099; 9803pp; English.
XX
XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
XX cancer-associated nucleic acid molecules (N) and proteins (P), where the
XX proteins are collectively known as colon cancer antigens. The colon
XX cancer antigens have cytostatic activity and can be used in gene therapy
XX and vaccine production. N and P may be used in the prevention, diagnosis
XX and treatment of diseases associated with inappropriate P expression. For
XX example, N and P may be used to treat disorders associated with decreased
XX expression by rectifying mutations or deletions in a patient's genome
XX that affect the activity of P by expressing inactive proteins or to
XX supplement the patients own production of P. Additionally, N may be used
XX to produce the colon cancer-associated Ps, by inserting the nucleic acids
```

CC into a host cell and culturing the cell to express the proteins. N and P
 CC can be used in the prevention, diagnosis and treatment of colorectal
 CC carcinomas and cancers. AAH37196 to AAH37204 and AAB7789 represent
 CC sequences used in the exemplification of the present invention. N.B.
 CC Pages 666 to 682 and page 7053 of the sequence listing were missing at
 CC time of publication, meaning no sequences are present for SEQ ID NO:1027
 CC to 1052, 7921 and 7922

XX Sequence 157 AA;

Query Match 76.6%; Score 791; DB 4; Length 157;
 Best Local Similarity 98.0%; Pred. No. 6,2e-69;
 Matches 150; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 44 SLGPAPELRHVLHNTLQOLQALRLAPALPPEPLFGEDPSLSATIGSLRELD 103
 :|||||
 DB 5 ALGPAPELRHVLHNTLQOLQALRLAPALPPEPLFGEDPSLSATIGSLRELD 64
 QY 104 TSMGTEPPNPVPTPLGQNEVPVPPDPVFLEALSSRYLGDGLDFFLDITDSAVEKEP 163
 65 TSMGTEPPNPVPTPLGQNEVPVPPDPVFLEALSSRYLGDGLDFFLDITDSAVEKEP 124
 QY 164 AAPPEPHNLFCAPGSWMNLDHIMEILIGS 196
 :|||||
 DB 125 AAPPEPHNLFCAPGSWMNLDHIMEILIGS 157

RESULT 6

ABO60075
 ID ABO60075 standard; protein; 142 AA.

XX ABO60075;

XX 29-JUL-2004 (first entry)

XX Human genome derived single exon protein #6309.

XX Human; gene expression; single exon probe; microarray;
 XX alternative splicing event; genomic alteration.

XX Homo sapiens.

XX US2003194704-A1.

XX 16-OCT-2003.

XX 03-APR-2002; 2002US-00029386.

XX 03-APR-2002; 2002US-00029386.

XX (PENN/) PENN S G.
 XX (RANK/) RANK D R.
 XX (HANZ/) HANZEL D K.

XX Penn SG, Rank DR, Hanzel DK;

XX MPI; 2004-119264/12.

XX New human genome-derived single exon nucleic acid probes useful for human
 XX gene expression analysis, for identifying or characterizing alternative
 XX splicing events, for assessing genomic alterations or as tools for
 XX surveying tissues.

XX Claim 45; SEQ ID NO 33709; 80bp; English.

XX The invention relates to a nucleic acid probe for measuring human gene
 XX expression, comprising any of the 27,400 fully defined nucleotide
 XX sequences in the specification, or their complements or fragments, and
 XX encoding at least 8 amino acids of any of the 688 amino acid sequences
 XX fully defined in the specification. The probe is a single exon probe that
 XX hybridizes under high stringency conditions to a nucleic acid molecule
 XX expressed in human cells or tissues. Also included are a spatially-
 XX addressable set of single exon nucleic acid probes for measuring human

CC gene expression (comprising a plurality of single exon nucleic acid
 CC probes cited above, where each of the plurality of probes is separately
 CC and addressably isolatable or amplifiable from the plurality), a single
 CC exon microarray for measuring human gene expression, a method of
 CC measuring human gene expression, a vector comprising the single exon
 CC probe cited above, an ORF-encoded peptide comprising at least 8
 CC contiguous amino acids of any of the above-mentioned amino acid
 CC sequences (optionally with conservative amino acid substitutions), an
 CC isolated antibody that binds specifically to a peptide cited above,
 CC a method of selling and/or licensing single exon probes or microarrays to
 CC a customer desiring to measure gene expression, a method of providing
 CC human gene expression data by subscription, and a computer-readable
 CC storage medium which contains a database having a plurality of records
 CC (each record including data on the expression of a single exon probe
 CC cited above. The probe, methods and apparatus are useful in gene
 CC expression analysis. The probes may be used as tools for surveying
 CC tissues to detect the presence of expressed messages that contain their
 CC specific exon, or in constructing genome-derived single exon microarrays.
 CC In addition, the probes are used in identifying and characterizing
 CC alternative splicing events, in detecting and characterizing gross
 CC alterations in the genomic locus that includes their exon, in assessing
 CC smaller genomic alterations, in priming the synthesis of nucleic acids,
 CC or in expressing the ORF-encoded peptide. The present sequence is a human
 CC single exon probe protein of the invention. Note: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from USPTO at
 CC seqdata.uspto.gov/sequence.html?docID=20030194704

XX Sequence 142 AA;

Query Match 62.5%; Score 645.5; DB 8; Length 142;
 Best Local Similarity 96.8%; Pred. No. 9e-55;
 Matches 122; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 71 LAPALPPEPLFGEDPSLSATIGSLRELDTSMDGTEPPNPVPTPLGQNEVPPOPD 130
 :|||||
 DB 18 LPPLPC-PPEPLFGEDPSLSATIGSLRELDTSMDGTEPPNPVPTPLGQNEVPPOPD 76
 QY 131 PVFLEALSSRYLGDGLDFFLDITDSAVEKEPARAPPEPHNLFCAPGSWMNLDHIM 190
 :|||||
 DB 77 PVFLEALSSRYLGDGLDFFLDITDSAVEKEPARAPPEPHNLFCAPGSWMNLDHIM 136
 QY 191 EILIGS 196
 :|||||
 DB 137 EILIGS 142

RESULT 7

ABO60074
 ID ABO60074 standard; protein; 111 AA.

XX ABO60074;

XX 29-JUL-2004 (first entry)

XX Human genome derived single exon protein #6308.

XX Human; gene expression; single exon probe; microarray;
 XX alternative splicing event; genomic alteration.

XX Homo sapiens.

XX US2003194704-A1.

XX 16-OCT-2003.

XX 03-APR-2002; 2002US-00029386.

XX 03-APR-2002; 2002US-00029386.

XX (PENN/) PENN S G.
 XX (RANK/) RANK D R.
 XX (HANZ/) HANZEL D K.

XX Penn SG, Rank DR, Hanzel DK;
 PI WPI; 2004-119264/12.
 XX
 XX
 PT New human genome-derived single exon nucleic acid probes useful for human
 PT gene expression analysis, for identifying or characterizing alternative
 PT splicing events, for assessing genomic alterations or as tools for
 PT surveying tissues.
 XX
 PS Claim 45; SEQ ID NO 33708; 80bp; English.
 XX
 XX The invention relates to a nucleic acid probe for measuring human gene
 CC expression, comprising any of the 27,400 fully defined nucleotide
 CC sequences in the specification, or their complements or fragments, and
 CC encoding at least 8 amino acids of any of the 688 amino acid sequences
 CC fully defined in the specification. The probe is a single exon probe that
 CC hybridises under high stringency conditions to a nucleic acid molecule
 CC expressed in human cells or tissues. Also included are a spatially-
 CC addressable set of single exon nucleic acid probes for measuring human
 CC gene expression (comprising a plurality of single exon nucleic acid
 CC probes cited above, where each of the plurality of probes is separately
 CC and addressably isolatable or amplifiable from the plurality), a single
 CC exon microarray for measuring human gene expression, a method of
 CC measuring human gene expression, a vector comprising the single exon
 CC probe cited above, an ORF-encoded peptide comprising at least 8
 CC contiguous amino acids of any of the above-mentioned amino acid
 CC sequences (optionally with conservative amino acid substitutions), an
 CC isolated antibody that binds specifically to a peptide cited above,
 CC methods of selling and/or licensing single exon probes or microarrays to
 CC a customer desiring to measure gene expression, a method of providing
 CC human gene expression data by subsequence, and a computer-readable
 CC storage medium which contains a database having a plurality of records
 CC (each record including data on the expression of a single exon probe
 CC cited above. The probe, methods and apparatus are useful in gene
 CC expression analysis. The probes may be used as tools for surveying
 CC tissues to detect the presence of expressed messages that contain their
 CC specific exon, or in constructing genome-derived single exon microarrays.
 CC In addition, the probes are used in identifying and characterising
 CC alternative splicing events, in detecting and characterising gross
 CC alterations in the genomic locus that includes their exon, in assessing
 CC smaller genomic alterations, in priming the synthesis of nucleic acids,
 CC or in expressing the ORF-encoded peptide. The present sequence is a human
 CC single exon probe protein of the invention. Note: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from USPTO at
 CC seqdata.uspto.gov/sequence.html?DocID=20030194704
 CC
 XX
 SQ Sequence 111 AA;
 Query Match 38.9%; Score 402; DB 8; Length 111;
 Best Local Similarity 89.9%; Pred. No. 4.1e-31;
 Matches 80; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
 QY 1 MEGGKRGKSDLEEEEREMWSPAGLQSYQALRLISLDKVRSGIPAPSLRRHVLTHN 60
 DB 3 MVGKRGKSDLEEEEREMWSPAGLQSYQALRLISLDKVRSGIPAPSLRRHVLTHN 62
 QY 61 TLQQLAALRLAPALPPEPLFLGEBDF 89
 DB 63 TLQQLAALRLAPALPPEPLPGRGPF 91
 QY
 DB
 RESULT 8
 AAY36004 standard; protein; 236 AA.
 XX
 AC AAY36004;
 XX
 DT 13-SEP-1999 (first entry)
 XX
 DB Extended human secreted protein sequence, SEQ ID NO. 389.
 XX

KW Secreted protein, human; cytokine; cellular proliferation; cell movement;
 KW cellular differentiation; immune system regulator; anti-inflammatory;
 KW haematopoiesis regulator; tissue growth regulator; tumour inhibitor;
 KW reproductive hormone regulator; chemotaxis; chemokinesis; gene therapy;
 KW genetic disease.
 XX
 OS Homo sapiens.
 XX
 PN WO9931236-A2.
 XX
 PD 24-JUN-1999.
 XX
 PF 17-DEC-1998; 98WO-IB002122.
 XX
 PR 17-DEC-1997; 97US-0069957P.
 XX
 PR 09-FEB-1998; 98US-0074121P.
 XX
 PR 13-APR-1998; 98US-0081563P.
 XX
 PR 10-AUG-1998; 98US-0096116P.
 XX
 PA (GSEST) GENSEST.
 XX
 PI Bouguetel L, Duclert A, Dumas Milne Edwards J;
 XX
 DR WPI; 1999-385906/32.
 XX
 DR N-PSDB; AAX97688.
 XX
 PT New isolated human secreted proteins.
 XX
 PS Claim 9; Page 330-331; 516pp; English.
 XX
 XX This sequence is encoded by an extended human secreted protein coding
 CC sequence of the invention. The secreted proteins can be used in treating
 CC or controlling a variety of human conditions. The secreted proteins may
 CC act as cytokines or may affect cellular proliferation or differentiation
 CC or may act as immune system regulators, haematopoiesis regulators, tissue
 CC growth regulators, regulators of reproductive hormones or cell movement
 CC or have chemotactic/chemokinetic, receptor/ligand, anti-inflammatory or
 CC tumour inhibition activity. The DNAs can be used in forensic procedures
 CC to identify individuals or in diagnostic procedures to identify
 CC individuals having genetic diseases resulting from abnormal expression of
 CC the genes corresponding to the extended cDNAs. They are also useful for
 CC constructing a high resolution map of the human chromosomes. They can
 CC also be used for gene therapy to control or treat genetic diseases
 CC
 XX
 SQ Sequence 236 AA;
 Query Match 20.9%; Score 216; DB 2; Length 236;
 Best Local Similarity 32.6%; Pred. No. 1.6e-12;
 Matches 76; Conservative 36; Mismatches 69; Indels 52; Gaps 13;
 QY 4 GLKRGKSDLEEEEREMWSPAGLQSYQALRLISLDKVRSGIPAPSLRRHVLTHN 51
 DB 5 GLKRGKSDLEEEEREMWSPAGLQSYQALRLISLDKVRSGIPAPSLRRHVLTHN 53
 QY 52 LRRHVLTHNTHQQLAALRLAPALPPEPLFLGEBDFSLSTIGSILRE 101
 DB 64 LRRHVLTHNTHQQLAALRLAPALPPEPLFLGEBDFSLSTIGSILRE 121
 QY 102 LDTSMGTEPPONPVPIGLQNEVPQPD-----PVFLAL-----SSRYLGDSGLDDPFL 152
 DB 122 L-SHTEGSLQAPQ-----LADGCPGRSISGXPPYLGLALGAPATCCLDNGEGFPE 175
 QY 153 DDTGSAVKE---PARAPPE-PPHLLFCARPSWE-----VNELDHIMELIIGS 196
 DB 176 DDTGSMYDNEIMAPASEGLKPGPD-----GPKKEAPELDEABLDVLMVLTGT 225
 QY
 DB
 RESULT 9
 ADP19312 standard; protein; 236 AA.
 XX
 AC ADP19312;
 XX

DT 26-AUG-2004 (first entry)
 XX Human secreted polypeptide #163.
 DE Human; secreted protein; genetic disease.
 XX Homo sapiens.
 OS US2004110939-A1.
 XX 10-JUN-2004.
 XX 15-OCT-2001; 2001US-00978360.
 XX 17-DEC-1998; 98WO-IB002122.
 XX 09-FEB-1999; 99WO-IB000282.
 XX 21-JUN-2000; 2000WO-IB000951.
 XX 15-SEP-2000; 2000US-00663600.
 XX (GIST) GENSET SA.
 PA Dumas Milne Edwards J, Bougueleret L, Jobert S, Clusel C;
 PI Duclert A;
 XX WPI; 2004-440404/41.
 XX N-PSDB; ADP18907.
 DR New isolated polynucleotide encoding secreted polypeptide, useful for
 PT gene therapy, or in diagnostic procedures to identify individuals having
 PT genetic diseases resulting from abnormal expression of the gene.
 XX Claim 2; SEQ ID NO 568; 113pp; English.
 PS The invention relates to human cDNA sequences that encode human secreted
 CC proteins. The invention also relates to an antibody that specifically
 CC binds to a polypeptide of the invention and a method of binding the
 CC polypeptide to an antibody. The polynucleotides are useful for expressing
 CC the entire secreted proteins which they encode and for distinguishing
 CC human tissues and cells from non-human tissues and cells, and for
 CC distinguishing between human tissues and cells that do or do not express
 CC the polynucleotides comprising the cDNAs. The polynucleotides and
 CC polypeptides are useful in forensic procedures or diagnostic procedures
 CC to identify individuals with genetic diseases resulting from abnormal
 CC expression of the genes corresponding to the cDNAs. The sequences are
 CC also useful in gene therapy to control or treat genetic diseases. This
 CC sequence represents a human secreted polypeptide of the invention. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification but was obtained in electronic format from USPRO at
 CC seqdata.uspto.gov/sequence.html.
 CC
 XX Sequence 236 AA;
 SQ

Query Match 20.9%; Score 216; DB 8; Length 236;
 Best Local Similarity 32.6%; Pred. No. 1.6e-12;
 Matches 76; Conservative 36; Mismatches 69; Indels 52; Gaps 13;

QY 4 GIKRKHSDLEEBE-----ERMWSPAGLOSYOQA-----LIRISLDKVRKSLGPRAPS 51
 DB 5 GIKRKHSDLEEBE-----ERMWSPAGLOSYOQA-----LIRISLDKVRKSLGPRAPS 63
 QY 52 LRRHVLHNTLQQLALRLAPALPPPEPL-----FLGEBDSLSATGSIIRB 101
 DB 64 LRRHVLHNTLQQLALRLAPALPPPEPL-----FLGEBDSLSATGSIIRB 121
 QY 102 LDTGMDGEPNPPTPIGLGNKVRPDP-----PFLRLAL-----SSRYLDSGSGLDXFL 152
 DB 122 L-SHIGLSQAPQP-----LADGGPGRSIGXPRYGLALDILGPRATGLDNGLSGLFE 175
 QY 153 DITSAVKEK---PARAPP-PRNLPAPGSGW-----WNELDHIMEIIGLS 196
 DB 176 DITSMYNEMLNAPASBGKRPED---GPGKEAPRLDRAELDYLMNVYOT 225

RESULT 10
 ID AAY44362 standard; protein; 236 AA.
 AC AAY44362;
 XX 14-MAR-2000 (first entry)
 XX Human cell cycle regulation protein-3.
 DE
 XX CECRP-3; cell cycle regulation protein-3; cell proliferation;
 KW cell proliferative disease; cancer; atherosclerosis; cirrhosis;
 KW hepatitis; psoriasis; immune system disorder; allergy; asthma;
 KW acquired immune deficiency syndrome; Crohn's disease; Blast method;
 KW rheumatoid arthritis; gene therapy; chromosomal mapping.
 XX Homo sapiens.
 OS
 XX

Key	Location/Qualifiers
Modified-site	44 /note= "Potential phosphorylation site"
Modified-site	60 /note= "Potential phosphorylation site"
Modified-site	73 /note= "Potential phosphorylation site"
Modified-site	98 /note= "Potential phosphorylation site"
Modified-site	117 /note= "Potential phosphorylation site"
Modified-site	123 /note= "Potential phosphorylation site"
Modified-site	180 /note= "Potential phosphorylation site"

W03964593-A2.
 16-DEC-1999.
 08-JUN-1999; 99WO-US012906.
 08-JUN-1998; 98US-0088695P.
 (INCYTE) INCYTE PHARM INC.
 Bandman O, Lal P, Tang YT, Corley NC, Guejler KJ, Baughn MR,
 Patterson C;
 WPI; 2000-105887/09.
 N-PSDB; AAZ29482.
 Novel regulatory proteins, for diagnosis, treatment and prevention of
 cell proliferative and immune system diseases.
 Claim 1; Page 70; 88pp; English.

The present sequence is a cell cycle regulation protein-3 (CECRP-3). Prints
 CC analytical method was used to identify this protein. CECRPs are
 CC activators of cell proliferation or inhibitors of cellular processes that
 CC modulate proliferation. They are used to treat or prevent cell
 CC proliferative diseases like cancers, atherosclerosis, cirrhosis,
 CC hepatitis, psoriasis, immune system disorders (e.g. acquired immune
 CC deficiency syndrome, allergy, asthma, Crohn's disease, rheumatoid
 CC arthritis). Antibodies are raised to screen for specific binding agents.
 CC The corresponding nucleic acid is used in gene therapy, chromosomal
 CC mapping and isolation of related sequences

Sequence 236 AA;
 SQ

Query Match 20.6%; Score 213; DB 3; Length 236;
 Best Local Similarity 32.6%; Pred. No. 3.2e-12;
 Matches 76; Conservative 36; Mismatches 69; Indels 52; Gaps 13;

QY 4 GIKRKHSDLEEBE-----ERMWSPAGLOSYOQA-----LIRISLDKVRKSLGPRAPS 51

```
Db      5  GKKRRKEEKEEKEPLAVDSW-WLDPGHAVVAQAPPAVAVSSLPDLISVLKTHSLQOSEPD 63
Qy      52  LRRHYLIHNTLQOOLAALRLAPALPPEPL-----FLGEDFSLATIGSILRE 101
Db      64  LRHLVLVNTTLRIRQAS--MAPAALPVPFSPPAAPSVADNLASSDALSMSLSLTD 121
Qy      102  LDTSMGTEPPQNPTVPLGLQNEVPP-----QDPVPELAL--SSRYLGDGSLDDPFL 152
Db      122  L-SHIEGLSQARQP-----LADEGPPGRSIGGAAPSLGALDILGPATGCLLDGLEGLE 175
Qy      153  DIDTSAVEKE--PARAPPEP-PHNLFCAPGSWE-----WNELDHIMEITLIS 196
Db      176  DIDTSMYDNEIWAAPASEGLKPGPED---GPGKEAPELDEARIDYIMDVAVGT 225

RESULT 11
AAM93724
ID      AAM93724 standard; protein; 236 AA.
XX
AC      AAM93724;
XX
DT      06-NOV-2001 (first entry)
XX
DE      Human polypeptide, SEQ ID NO: 3677.
XX
KW      Human; full length cDNA; cDNA synthesis; oligo-capping.
XX
OS      Homo sapiens.
XX
PN      EP1130094-A2.
XX
PD      05-SEP-2001.
XX
PF      07-JUL-2000; 2000EP-00114089.
XX
PR      08-JUL-1999; 99JP-00194486.
XX
PR      11-JAN-2000; 2000JP-00118774.
XX
PR      02-MAY-2000; 2000JP-00183765.
XX
PA      (HELIX) HELIX RES INST.
XX
PI      Ota T, Mishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y,
XX      Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX      MPI; 2001-524255/58.
XX      N-PSDB; AAK94674.
XX
PT      830 Primers useful for synthesizing full length cDNA clones and their use
XX      in genetic manipulation.
XX
PS      Claim 8; SEQ ID NO 3677; 1380bp + Sequence listing; English.
XX
CC      The invention relates to primers for synthesizing full length cDNA
XX      clones. 830 cDNA molecules encoding a human protein have been isolated
XX      and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
XX      been determined. Primers for synthesizing the full length cDNA are useful
XX      for clarifying the function of the protein encoded by the cDNA. The full
XX      length clones were obtained by construction of full length enriched cDNA
XX      libraries that were synthesized by the oligo-capping method. The primers
XX      enable the production of the full length cDNA easily without any special
XX      method. The present sequence is a polypeptide encoded by a full length
XX      human cDNA of the invention. Note: The sequence data for this patent did
XX      not form part of the printed specification, but was obtained in CD-ROM
XX      format directly from EPO
XX
SQ      Sequence 236 AA;

Query Match      20.6%; Score 213; DB 4; Length 236;
Best Local Similarity 32.6%; Pred. No. 3, 2e-12;
Matches 76; Conservative 36; Mismatches 69; Indels 52; Gaps 13;

4  GKKRKHSLDEE---ERWEWSPALQSQYQA-----LLRISLDKVGKSLGPAPAS 51
```

```
Db      5  GKKRKEEKEEKEPLAVDSW-WLDPGHAVVAQAPPAVAVSSLPDLISVLKTHSLQOSEPD 63
Qy      52  LRRHYLIHNTLQOOLAALRLAPALPPEPL-----FLGEDFSLATIGSILRE 101
Db      64  LRHLVLVNTTLRIRQAS--MAPAALPVPFSPPAAPSVADNLASSDALSMSLSLTD 121
Qy      102  LDTSMGTEPPQNPTVPLGLQNEVPP-----QDPVPELAL--SSRYLGDGSLDDPFL 152
Db      122  L-SHIEGLSQARQP-----LADEGPPGRSIGGAAPSLGALDILGPATGCLLDGLEGLE 175
Qy      153  DIDTSAVEKE--PARAPPEP-PHNLFCAPGSWE-----WNELDHIMEITLIS 196
Db      176  DIDTSMYDNEIWAAPASEGLKPGPED---GPGKEAPELDEARIDYIMDVAVGT 225

RESULT 12
AAG89292
ID      AAG89292 standard; protein; 236 AA.
XX
AC      AAG89292;
XX
DT      11-SEP-2001 (first entry)
XX
DE      Human secreted protein, SEQ ID NO: 412.
XX
KW      Human; secreted protein; gene therapy; vaccine; treatment; diagnosis;
XX      GENSET.
XX
OS      Homo sapiens.
XX
PN      WO200142451-A2.
XX
PD      14-JUN-2001.
XX
PF      07-DEC-2000; 2000WO-IB001938.
XX
PR      08-DEC-1999; 99US-0169629P.
XX
PR      06-MAR-2000; 2000US-0187470P.
XX
PA      (GENSET) GENSET.
XX
PI      Dunas Milne Edwards J, Bougueleret L, Jobert S;
XX      MPI; 2001-367870/38.
XX      N-PSDB; AAK64895.
XX
PT      Full length GENSET human nucleic acids encoding potentially secreted
XX      proteins, useful in gene therapy and vaccination against a variety of
XX      diseases, and for diagnosis of those diseases.
XX
PS      Claim 21; Page 889-890; 921bp; English.
XX
CC      The invention relates to full length GENSET human nucleic acids encoding
XX      potentially secreted proteins. The nucleic acids and the polypeptides
XX      they encode may be used in the prevention, treatment and diagnosis of
XX      diseases associated with inappropriate GENSET gene expression. For
XX      example, they be used to treat disorders associated with decreased GENSET
XX      gene expression by rectifying mutations or deletions in a patient's
XX      genome that affect the activity of GENSET or by supplementing the
XX      patients own production of GENSET polypeptides. Conversely, antisense
XX      nucleic acid molecules may be administered to down regulate GENSET
XX      expression by binding with the cells' own genes and preventing their
XX      expression. The sense and antisense nucleic acids may also be used as DNA
XX      probes in diagnostic assays to detect and quantitate the presence of
XX      similar nucleic acid sequences in samples, and hence to determine which
XX      patients may be in need of restorative therapy. The GENSET polypeptides
XX      may be used as antigens in the production of antibodies and in assays to
XX      identify modulators (agonists and antagonists) of GENSET polypeptide
XX      expression and activity. The present sequence is a GENSET polypeptide of
XX      the invention
XX
SQ      Sequence 236 AA;
```


CC they are expressed in, such as: antiinflammatory; antirheumatic;
 CC antiarthritic; immunosuppressive; antibacterial; endocrine; cardiac;
 CC central nervous system; virucide; anti-HIV; fungicide; antimutagen;
 CC cardiovascular; antianemic; antiagregant; haemostatic; vulnerary;
 CC antidiuretic; osteopathic; dermatological; antiallergic; antiaesthetic;
 CC antidiabetic; cyrostatic; neuroprotective; antidepressant; nootropic;
 CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides
 CC encoding them can be used in gene therapy, antisense therapy and vaccine
 CC production. The proteins and polynucleotides are useful for screening for
 CC agonists or antagonists of a protein and for the treatment and diagnosis
 CC of disorders associated with the activity of a protein e.g. inflammation,
 CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
 CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
 CC infections, autoimmunity, genetic diseases, haematopoietic disorders,
 CC anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers,
 CC osteoporosis, severe combined immunodeficiency, eczema, allergic
 CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
 CC Alzheimer's disease, Parkinson's disease, neurodegenerative and
 CC neurological disorders

CC Sequence 278 AA;

Query Match 20.6%; Score 213; DB 4; Length 278;
 Best Local Similarity 32.6%; Pred. No. 3.9e-12;
 Matches 76; Conservative 36; Mismatches 69; Indels 52; Gaps 13;

QY 4 GLKRGHSDLEEE-----ERMESWSPAGLOSVOOA-----LIRISLDKVGQSLGPRAPS 51
 DB 47 GLKRGREBEKEPLAVDSW-WLDPGHAAVAQAPPAVAVSSSLFDLSVTKLHNSLQOSEPD 105
 QY 52 LRRHYLINTIQQALRLAPALPPEPL-----FLGEDPSLSATTGSIIR 101
 DB 106 LRLHLVLTNLTIRIQAS--MAPAALPVPSPPAAPSAVDNLASDAAISMSASLLED 163
 QY 102 LDTSMGTEPPONPTPTGIGQNEVPP-----QDPVPLEAL--SSRYLGDSGIDDFPL 152
 DB 164 L-SHIEGISOAPQ-----LADGPPGRSIGGAAPSLGALDLGPAATCLDDGLEGF 217
 QY 153 DIDTSAVKE---PARAPPEP-PHNLFCAPGSWE-----WNELDHIMEIIGS 196
 DB 218 DIDTSMYNEIMAPASBGKPEPD--GPGKEAPBLDEABLDYLMQVLTGT 267

RESULT 15

AA31829
 ID AA31829 standard; protein; 236 AA.

XX AC AA31829;

XX DT 06-DEC-1999 (first entry)

XX DE Human adult blood secreted protein g21_1.

XX KW Secreted protein; g21_1; human; therapy; diagnosis; vaccine; blood.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Misc-difference 1 /note= "encoded by AWG"

FT Peptide 19..21 /note= "signal peptide"

FT Protein 32..236 /note= "mature protein"

FT Domain 40 /note= "a putative transmembrane domain is centered
 around this residue"

FT Peptide 67..79 /note= "alternative signal peptide"

FT Protein 80..236 /note= "alternative mature protein"

FT Domain 80 /note= "a putative transmembrane domain is centered
 around this residue"

FT Misc-difference 137 around this residue"
 FT /note= "encoded by GAS"
 FT Domain 150 /note= "a putative transmembrane domain is centered
 around this residue"

FN WO9947555-A1.

PD 23-SEP-1999.

PF 18-MAR-1999; 99WO-US005939.

PR 20-MAR-1998; 98US-0078803P.

PR 17-MAR-1999; 99US-00078803.

PA (GENY) GENETICS INST INC.

PI Jacobs K, Mccoy JM, Lavallic ER, Collins-Racie LA, Evans C;

PI Merberg D, Treacy M, Agoetino MJ, Steinger RJ;

DR N-PSDB; AA219894.

PT New polynucleotides derived from murine fetal cell cDNA libraries,

PT potentially used as, e.g. vaccines.

PS Claim 13(a); Page 94; 107pp; English.

CC This is the predicted amino acid sequence of a novel human secreted
 CC protein, g21_1, as deduced from an isolated adult blood cDNA clone (see
 CC AA219894). The invention provides new human secreted proteins (see
 CC AA31828-38) and polynucleotides (see AA219893-901) isolated from foetal
 CC cell, adult blood, adult brain and foetal kidney cDNA libraries. They are
 CC predicted to have biological activities which would make them suitable
 CC for treating, preventing or ameliorating medical conditions in humans and
 CC animals, although no supporting data are given. Suggested activities
 CC include nutritional, cytokine, tissue growth, cell proliferation and
 CC differentiation, immunostimulant (e.g. as vaccine), immunosuppressive,
 CC haematopoiesis regulating, activin or inhibin, chemotactic or
 CC chemokinetic, haemostatic or thrombolytic, receptor/ligand activity,
 CC antiinflammatory, cadherin or tumour invasion suppressor, and tumour
 CC inhibition activities

CC Sequence 236 AA;

Query Match 20.4%; Score 211; DB 2; Length 236;
 Best Local Similarity 32.6%; Pred. No. 5e-12;
 Matches 76; Conservative 35; Mismatches 70; Indels 52; Gaps 13;

QY 4 GLKRGHSDLEEE-----ERMESWSPAGLOSVOOA-----LIRISLDKVGQSLGPRAPS 51
 DB 5 GLKRGREBEKEPLAVDSW-WLDPGHAAVAQAPPAVAVSSSLFDLSVTKLHNSLQOSEPD 63
 QY 52 LRRHYLINTIQQALRLAPALPPEPL-----FLGEDPSLSATTGSIIR 101
 DB 64 LRLHLVLTNLTIRIQAS--MAPAALPVPSPPAAPSAVDNLASDAAISMSASLLED 121
 QY 102 LDTSMGTEPPONPTPTGIGQNEVPP-----QDPVPLEAL--SSRYLGDSGIDDFPL 152
 DB 122 L-SHIEGISOAPQ-----LAXEGPPGRSIGGAAPSLGALDLGPAATCLDDGLEGF 175
 QY 153 DIDTSAVKE---PARAPPEP-PHNLFCAPGSWE-----WNELDHIMEIIGS 196
 DB 176 DIDTSMYNEIMAPASBGKPEPD--GPGKEAPBLDEABLDYLMQVLTGT 225

Search completed: February 6, 2006, 09:56:50
 Job time : 203 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 6, 2006, 09:57:06 / Search time 43 Seconds
(without alignments)
438.569 Million cell updates/sec

Title: US-10-069-386a-2

Perfect score: 1033

Sequence: 1 MEGGLKRRKHSDLSEERERME.....APGSWMNLDHIMEIILGS 196

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	101.5	9.8	728	2 HS9435	phosphoinositide-3
2	97	9.4	803	2 T40514	Chaperonin hsp78p
3	97	9.4	1201	2 G86441	unknown protein (1
4	96	9.3	817	2 S51342	verprolin - yeast
5	95.5	9.2	446	2 A43029	transcription fact
6	93.5	9.1	1734	2 A54602	microtubule-associ
7	93	9.0	433	1 FOLJH2	gag polyprotein -
8	91.5	8.9	1621	2 A82255	hypothetical prote
9	90.5	8.8	804	2 AG0565	probable membrane
10	88	8.5	651	2 T42644	hypothetical prote
11	88	8.5	716	2 T26998	hypothetical prote
12	88	8.5	1676	2 A55508	amuleste primary
13	86.5	8.4	536	2 A34596	transcription fact
14	86.5	8.4	1494	2 T14355	protein-tyrosine-p
15	86	8.3	475	2 B37761	ntrA protein - Thi
16	85.5	8.3	485	2 T37550	hypothetical colle
17	85.5	8.3	825	2 E75508	conserved hypothet
18	85.5	8.3	1544	2 E59431	phosphoinositide-b
19	85	8.2	715	2 D85087	hypothetical prote
20	85	8.2	1217	2 T00270	hypothetical prote
21	84.5	8.2	366	1 SAVLMD	large surface anti
22	84.5	8.2	477	2 T46304	hypothetical prote
23	84.5	8.2	723	2 B38749	3-phosphatidylinos
24	84.5	8.2	908	2 T15057	hypothetical prote
25	84.5	8.2	982	2 T43676	humback-related
26	84.5	8.2	1262	2 T13353	protein-tyrosine-p
27	84	8.1	289	2 D70452	leucine-tRNA ligase
28	84	8.1	954	2 D16114	co-repressor prote
29	84	8.1	1015	2 JCS062	phogrin precursor

30	84	8.1	1026	2 T28968	hypothetical prote
31	84	8.1	1234	2 T00363	hypothetical prote
32	84	8.1	1819	2 T32008	hypothetical prote
33	83.5	8.1	331	2 S09800	hypothetical prote
34	83.5	8.1	601	2 AG0066	probable AMP-bindi
35	83.5	8.1	804	2 A85549	probable oxidoredu
36	83.5	8.1	804	2 G90698	hypothetical prote
37	83	8.0	346	2 E72672	DNA-binding protei
38	83	8.0	737	2 S28030	icht protein - Ink
39	83	8.0	1353	2 T00249	structural protein
40	82.5	8.0	555	2 T30349	protein-tyrosine k
41	82.5	8.0	1130	1 TVH0A	MHC class III hist
42	82.5	8.0	1870	2 S37671	MHC class III hist
43	82.5	8.0	1872	2 S36152	MHC class III hist
44	82.5	8.0	2142	2 B35098	MHC class III hist
45	82	7.9	450	2 B47265	tailless (tll) pro

ALIGNMENTS

RESULT 1
HS9435
phosphoinositide-3-kinase regulatory beta chain [imported] - human
C:Species: Homo sapiens (man)
C:Date: 03-Jun-2002 #sequence_revision 03-Jun-2002 #text_change 09-Jul-2004
C:Accession: HS9435; A59436
R:Volinia S; Patrachini P; Otsu M; Hiles I; Gout I; Calzolari E; Bernardi F; Rooke L; 'Oncogene 7', 789-793, 1992
A:Title: Chromosomal localization of human p85 alpha, a subunit of phosphatidylinositol
A:Reference number: HS9435
A:Accession: HS9435
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-728 <VOL>
A:Cross-references: UNIPROT:000459; UNIPARC:UPI000013106C; GB:NP_005018; PID:g4826908; PID:NP_005018.1
R:Janssen, J.W.; Schleithoff, L.; Barttram, C.R.; Schulz, A.S.
Oncogene 16, 1767-1772, 1998
A:Title: An oncogenic fusion product of the phosphatidylinositol 3-kinase p85beta subunit
A:Reference number: A59436; MUID:98241181; PMID:9582025
A:Accession: A59436
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-728 <JAN>
A:Cross-references: UNIPARC:UPI000013106C; GB:NP_005018; PID:g4826908; PID:NP_005018.1

Query Match 9.8%; Score 101.5; DB 2; Length 728;
Best Local Similarity 23.9%; Pred. No. 1;
Matches 47; Conservative 25; Mismatches 70; Indels 55; Gaps 9;

QY 20 EWSPA-----GLOSVOAL-----LRISLDKYQRLGPRAPRLRHVLIH 59
DB 164 QMDTALADGKISFLAPLVTPEASABARRARLRAGPGLPEPTPLRHPLALTR 223
QY 60 NTLOQLALRLAPALPPEPLFGSEBDSISATIGSLRLRLDTSMDGTBPPQNVPL 119
DB 224 FILQHLGRVARRAP-----LGPVRALGATFGPL-----LRAPPSPSP-PPG 267
QY 120 GLQNEVPPQPD--PVFLALSSRYLGDGLDFLIDITSAVEKAPRAPPP----- 170
DB 268 GAPDSBSPSPDPALVLVEKLQENHLEQ-----EVAPRALPPRPKAPRYLANCG 320
QY 171 -PHNIFCAPGSWMNML 186
DB 321 SPPSL--QDAEWYMGDI 335

RESULT 2
T40514
Chaperonin hsp78p - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T40514

Db 634 ASLADVLPSSNLKRPSPVAAAPLPTFSAPSLPQGSVSTISIPSPV-APTLSVR-- 690
 QY 143 GDSGLDPLDIDTSAVEKSPAPAPPEP 171
 Db 691 -----TETESISKNPTKSPPPP 708

RESULT 5

A42029
 transcription factor B3 - mouse
 C:Species: Mus musculus (house mouse)
 C>Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
 C:Accession: A42029
 R:Roman, C.; Matera, A.G.; Cooper, C.; Arlandi, S.; Blain, S.; Ward, D.C.; Calame, K.
 Mol. Cell. Biol. 12, 817-827, 1992
 A>Title: mTEB3, an X-linked transcriptional activator containing basic helix-loop-helix
 A:Reference number: A42029; MUID:92123207; PMID:1732746
 A:Accession: A42029
 A:Status: Preliminary
 A:Molecule type: mRNA
 A:Residues: 1-446 <RNA>
 A:Cross-references: UNIPROT:Q64092; UNIPARC:UPI0000029662; GB:S76673; NID:9243439; PIND:
 A>Note: sequence extracted from NCBI backbone (NCBIN:76673, NCBI:P:76674)

Query Match 9.2%; Score 95.5; DB 2; Length 446;
 Best Local Similarity 27.3%; Pred. No. 1.7;
 Matches 48; Conservative 24; Mismatches 61; Indels 43; Gaps 8;

QY 12 LEEBERMWSVAGLQSVQALRLISLDKVRSLGPRASLRHVLIHNTLQOLQALRL 71
 Db 272 LKQEQOR-----SKDLESRQR-----SLQANSLQLRIQL-----ELQALING 311
 QY 72 AAPALPPEPLPLGSEDFSLTIGSLRLDLTSMGTPPPQNPVPLGLQNEVPPQPP 131
 Db 312 LP---VPPVPLGLSLTSSVSDSLKP--EQLDIEERGRSTTFHVSQGPAPQAPQPPA 366
 QY 132 VFLMAL-----SSRYLGDGLDPLDITSAVEK-----PARAPPEP 170
 Db 367 PPSDALLDHPFSDHLGLG-DPFLGLIEDIMEERGVGSLGALSLPLRAASP 421

RESULT 6

A54602
 microtubule-associated serine/threonine protein kinase MA2705 - mouse
 C:Species: Mus musculus (house mouse)
 C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 31-Dec-2004
 C:Accession: A54602
 R:Walden, P.D.; Cowan, N.J.
 Mol. Cell. Biol. 13, 7625-7635, 1993
 A>Title: A novel 205-kilodalton testis-specific serine/threonine protein kinase associat
 A:Reference number: A54602; MUID:94067123; PMID:8246979
 A:Accession: A54602
 A:Status: Preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1734 <MAL>
 A:Cross-references: UNIPROT:Q60592; UNIPARC:UPI0000027866; GB:U02313; NID:9406057; PIND:
 F:451-726/Domain: protein kinase ATP-binding motif
 F:459-467/Region: protein kinase ATP-binding motif

Query Match 9.1%; Score 93.5; DB 2; Length 1734;
 Best Local Similarity 26.3%; Pred. No. 15;
 Matches 51; Conservative 23; Mismatches 71; Indels 49; Gaps 8;

QY 25 GLQSVQAL-----LRISLDKVRSLGPRASLRHVLIHNTLQOLQALRLAPAPPEP 79
 Db 1308 GDSGLDPLDIDTSAVEKSPAPAPPEP-----VQSKLAAALAAAR 1357
 QY 80 EPLPLGSEDFSLTIGSLRLDLTSMGTPPPQNPVPLGLQNEVPP-----PPDPVFLA 136
 Db 1358 KKL-APSRKSLDPLGELKEL-----TPREASPLFVGVTSVLSGKGPLGKGVLP 1410

QY 137 LSSRYLG-----DSGLDPLDIDTSAVEKSPAPAP-PEPPHN 173
 Db 1411 APSRALGTRQDRAARRRSLQKQALIRVDSSEDDTDESPNSCATQSPRLSPHEASHN 1470
 QY 174 LFCARSWENMELD 187
 Db 1471 LL-PRGSGEGTEED 1483

RESULT 7

POLJH2
 gag polyprotein - human T-cell lymphotropic virus type 2
 N:Alternate names: core polyprotein
 N:Contains: core protein p12; core protein p15; core protein p24
 C:Species: human T-cell lymphotropic virus type 2, HTLV-2
 A>Note: host Homo sapiens (man)
 C>Date: 28-Feb-1986 #sequence_revision 28-Feb-1986 #text_change 09-Jul-2004
 C:Accession: A03944
 R:Shimotohno, K.; Takahashi, Y.; Shimizu, N.; Gojobori, T.; Golde, D.W.; Chen, I.S.Y.; I
 Proc. Natl. Acad. Sci. U.S.A. 82, 3101-3105, 1985
 A>Title: Complete nucleotide sequence of an infectious clone of human T-cell leukemia vi
 A:Reference number: A94042; MUID:85216449; PMID:2582407
 A:Accession: A03944
 A:Molecule type: DNA
 A:Residues: 1-433 <SH1>
 A:Cross-references: UNIPROT:P03346; UNIPARC:UPI000012B002; GB:M10060; NID:9329559; PIND
 A>Note: the authors translated the codon TAC for residues 197 and 249 as Thr
 C:Genetics:
 A:Gene: gag
 C:Superfamily: mammalian retrovirus gag polyprotein II
 C:Keywords: core protein; polyprotein
 F:1-136/Product: core protein p15 #status predicted <P15>
 F:137-214/Product: core protein p24 #status predicted <P24>
 F:215-433/Product: core protein p12 #status predicted <P12>

Query Match 9.0%; Score 93; DB 1; Length 433;
 Best Local Similarity 23.1%; Pred. No. 2.7;
 Matches 40; Conservative 24; Mismatches 67; Indels 42; Gaps 7;

QY 47 PRAPSLRHHVLIHNTLQOLQALRLAPAP-----LPPEPLFGSEDFSLA 93
 Db 13 PRAP---RGLSTHMLNPLQAAVRLQPPRSPDFQQLRRFLKALTKTPINLPIDYSLA 69
 QY 94 TI-----GSILRLDIDTSMGTPPPQNPVPLG---LQNEVPPQDPVFLA-LSSRYLG 143
 Db 70 SLIPKGYGRVVEIINILVKQVSSAPAAVPTPTCTTTPPPPPSPAHVPPV-- 127
 QY 144 DSGLDPLDIDTSAVEKSPAPAPPEPPHNLFCARSWENMELDHIMEIILGS 196
 Db 128 -----VPTTTCFPIHPGAP-----SAHRPMQMDLQAIKQVSSS 166

RESULT 8

A82255
 hypothetical protein VC0998 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
 C:Species: Vibrio cholerae
 C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
 C:Accession: A82255
 R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.;
 chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, P
 1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
 Nature 406, 477-483, 2000
 A>Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A:Reference number: A82035; MUID:20406833; PMID:10952301
 A:Accession: A82255
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-1621 <HRI>
 A:Cross-references: UNIPROT:Q9KTA5; UNIPARC:UPI00000C2BAC; GB:AB004181; GB:AB003852; NID
 A:Experimental source: serogroup O1; strain N16961; biotype El Tor
 C:Genetics:
 A:Gene: VC0998
 A:Map position: 1

A;Residues: 1-1676 <FIS>
A;Cross-references: UNIPROT:Q00083; UNIPARC:UPI0000125GB; GB:X82289; MID:G683499; PID:9
C;Genetics:
A;Gene: apaA
A;Introns: 149/3

Query Match 8.5%; Score 88; DB 2; Length 1676;
Best Local Similarity 20.5%; Pred. No. 41;
Matches 32; Conservative 19; Mismatches 51; Indels 54; Gaps 6;

QY 36 ISLDKVRSGSLPPAPSLRRHVLHNTLQQLAALRLAPALPPPEPLFGSEEDSLSTI 95
DB 837 VSFEBTE-PVAPSPPELRTAFVGSSTTEPVAPVPEVALSP-----I 880
QY 96 GSIRELDTSMGTEPPQNPVTPPLGLQNEVPFPQDPVFL----- 134
DB 881 SS-----QTTQPEPVPV-----APPEPFPVPEMPSQILVEDTLPIATLP 924
QY 135 EALSSRYLGDGSLDFFLIDITSAVEKPPARAPPEP 170
DB 925 EPAPERVAEQTSTTDIADVSVAISSQTE-PVER 959

RESULT 13

A34596
transcription factor B3 - human (fragment)

C;Species: Homo sapiens (man)

C;Date: 06-Jul-1990 #sequence_revision 09-Oct-1992 #text_change 09-Jul-2004

C;Accession: A34596; S10379

R;Beckmann, H.; Su, L.K.; Kadesch, T.

Genes Dev. 4, 167-179, 1990

A;Title: TFE3: a helix-loop-helix protein that activates transcription through the immun

A;Reference number: A34596; MUID:90249724; PMID:2338243

A;Accession: A34596

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-536 <BEC>

A;Cross-references: UNIPROT:P19532; UNIPARC:UPI000016B0B0; EMBL:X51330; NID:937061; PIDN

A;Note: the authors translated the codon ACC for residue 433 as Ser, and GAG for residue

C;Genetics:

A;Gene: GDB:TFE3

A;Cross-references: GDB:125870; OMIM:314310

A;Map position: Xp11.23-Xp11.22

C;Keywords: DNA binding; transcription factor

Query Match 8.4%; Score 86.5; DB 2; Length 536;
Best Local Similarity 25.4%; Pred. No. 13;
Matches 47; Conservative 24; Mismatches 55; Indels 59; Gaps 10;

QY 12 LEEERERWESPAGLOSVOQLRLISLDKVRSGSLPPAPSLRRHVLHNTLQQLAALRL 71
DB 192 LKQKQQR---SKDLERQR---SLQANRSLQRLQEL-----ELQAQIHG 231
QY 72 APAPALPPPEPLFGSEEDSLSTIGSLIRE-LDTSMGTE-----PPQNPVTPPLG 120
DB 232 LPVGTG-----GLSLATSTSDSLKPEQDIEEBRPPGARFTHVGGGAQN----- 279
QY 121 LQNEVPFP--DPVFLALSSRYLGDGSLDFFLIDITSAVEK-----PAR 165
DB 280 APHQQPPAPSPDALDLHFPDNLGDLG-DFPHGLGLIEDILMEEEGVVGLSGALSLR 338
QY 166 APPP 170
DB 339 AASDP 343

RESULT 14

T1435

protein-tyrosine-phosphatase (EC 3.1.3.48) TD14 - rat

C;Species: Rattus norvegicus (Norway rat)

C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004

C;Accession: T14355

R;Cao, L.; Zhang, L.; Ruiz-Lozano, P.; Yang, Q.; Chien, K.R.; Graham, R.M.; Zhou, M.

J. Biol. Chem. 273, 21077-21083, 1998
A;Title: A novel putative protein-tyrosine phosphatase contains a BCR1-like domain and s
A;Reference number: 218004; MUID:98361981; PMID:9694860
A;Accession: T14355
A;Status: preliminary; translated from GB/EMBL/DBJ

A;Residues: 1-1494 <CAO>
A;Cross-references: UNIPROT:Q08902; UNIPARC:UPI000005DAD; EMBL:AF077000; NID:93598973;
A;Experimental source: brain
C;Genetics:

A;Gene: PTP-TD14

C;Function: may be involved in regulating Ha-rae-dependent cell growth

C;Keywords: phosphoric monoester hydrolase

Query Match 8.4%; Score 86.5; DB 2; Length 1494;
Best Local Similarity 29.1%; Pred. No. 47;
Matches 37; Conservative 19; Mismatches 40; Indels 31; Gaps 9;

QY 62 LQQLAALRLAPALPPPEPLFGSEEDSLSTIGSLIRE-----LDTSMGTEPPQNP 115
DB 625 LMQPRAPVPMAGVLPAPVTSB-----LGLVPRSSPQHGIVSSPVAGVGPQ-P 675
QY 116 VTPVGLQNEVPQ--PDVFLALSSRYLGDGSLDFFLIDITSAVEKPPARAPPEP 171
DB 676 I--VGLPSAPPQPSGPE---LAMVDR-PATTVDSVOAPISSHMLRPPAPAPQP-P 727
QY 172 HNLFCAP 178
DB 728 ---CFP 730

RESULT 15

B37761

ntrA protein - Thioobacillus ferrooxidans

C;Species: Thioobacillus ferrooxidans

C;Date: 31-May-1991 #sequence_revision 31-May-1991 #text_change 09-Jul-2004

C;Accession: B37761

R;Berger, D.K.; Woods, D.R.; Rawlings, D.E.

J. Bacteriol. 172, 4399-4406, 1990

A;Title: Complementatation of Escherichia coli sigma(54) (ntrA)-dependent formate hydrogen

A;Reference number: A37761; MUID:90330545; PMID:2198257

A;Accession: B37761

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-475 <BER>

A;Cross-references: UNIPROT:P24695; UNIPARC:UPI0000134582; GB:M58480; GB:M33831; NID:915

C;Superfamily: Pseudomonas transcription initiation factor sigma

C;Keywords: DNA binding; transcription regulation

Query Match 8.3%; Score 86; DB 2; Length 475;
Best Local Similarity 21.7%; Pred. No. 12;
Matches 40; Conservative 31; Mismatches 75; Indels 38; Gaps 6;

QY 1 MEGGLKRGHSDLEEERWESPAGLOSVOQLRLISLDKVRSGSLPPAPSLRRHVLHNTLQQLAALRL 53
DB 307 MAGGKDAHKHYIQQLWEARFISLSGRDITLKVAIYERKQDFRANPESRPPVWL 366
QY 54 RHVLHNTLQQLAALRLAPALPPPEPLFGSEEDSLSTIG-----SLIRELDT 104
DB 367 RH--IADAVEHEESTVSRVTVQKIMTFRGLYEFKYPSSHVGSDGSASATIRALLI 424
QY 105 SMDGTEPPQNPVTPPLGLQNEVPFPQDPVFLALSSRYLGDGSLDFFLIDITSAVEKPPA 164
DB 425 KMTQAEQAQHPLS-----DAETARVLADGIG--IARRTVAKYREAA 464
QY 165 RAPP 168
DB 465 NVPP 468

Search completed: February 6, 2006, 10:01:57

Job time : 46 secs

Protein Sequence Searches - February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension **.rup**) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (UniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: February 6, 2006, 09:53:36 ; Search time 253 Seconds

(without alignments)
546.575 Million cell updates/sec

Title: US-10-069-386a-2

Perfect score: 1033
Sequence: 1 MEGLIKRRKSHDLSEERWE.....AFGSWENLHIMEIILGS 196

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :
1: uniprotc_05.80:*
2: uniprotc_sprotl:*
3: uniprotc_trembl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1026	99.3	196	1 SRTD3_HUMAN	Q9UW9 homo sapien
2	871	84.3	197	1 SRTD3_MOUSE	Q9UW9 mus musculi
3	858	83.1	197	2 OSBK27_RAT	Q5BK27 rattus norv
4	217	21.0	236	1 SRTD1_MOUSE	Q9J110 mus musculi
5	216.5	21.0	244	2 OSBK00_BRARE	Q5BEQ0 brachydanto
6	213	20.6	236	1 SRTD1_HUMAN	Q9UW2 homo sapien
7	213	20.6	236	2 OS3GC0_HUMAN	Q53GC0 homo sapien
8	208	20.1	236	2 OSBP771_RAT	Q6P771 rattus norv
9	166	16.1	244	2 Q9DC22_MOUSE	Q9DC22 mus musculi
10	163.5	15.8	244	2 Q4T2M4_TETNG	Q4T2M4 tetraodon n
11	145	14.0	275	2 Q4S6F1_TETNG	Q4S6F1 tetraodon n
12	144.5	14.0	229	2 OSRIK3_BRARE	Q5RIK3 brachydanto
13	140.5	13.6	309	1 SRTD2_MOUSE	Q9J195 mus musculi
14	140.5	13.6	309	2 OSSES5_MOUSE	Q5SES5 mus musculi
15	136.5	13.2	248	2 Q6CM81_XENLA	Q6CM81 xenopus lae
16	136	13.2	237	1 CDCA4_MOUSE	Q9CWM2 mus musculi
17	134.5	13.0	315	2 Q4V7A7_RAT	Q4V7A7 rattus norv
18	133	12.9	361	2 Q6NMD9_BRARE	Q6NMD9 brachydanto
19	133	12.9	361	2 Q7ZZ27_BRARE	Q7ZZ27 brachydanto
20	132	12.8	316	2 OSZJES_CHICK	Q5ZJES gallus gall
21	130.5	12.6	286	2 OSRIK2_BRARE	Q5RIK2 brachydanto
22	125	12.1	314	1 SRTD2_HUMAN	Q14140 homo sapien
23	125	12.1	314	2 Q53TS2_HUMAN	Q53TS2 homo sapien
24	125	12.1	314	2 Q5RDH4_PONPY	Q5RDH4 pongo pygma
25	121	11.7	364	2 Q4S6G5_TETNG	Q4S6G5 tetraodon n
26	115	11.1	1194	2 Q76M68_RAT	Q76M68 rattus norv
27	113.5	11.0	911	2 Q80TJ8_MOUSE	Q80TJ8 mus musculi
28	113	10.9	249	2 Q5ZL47_CHICK	Q5ZL47 gallus gall
29	112.5	10.9	246	2 Q6DFP1_XENLA	Q6DFP1 xenopus lae
30	105.5	10.2	757	1 C8P68_HUMAN	Q76J32 homo sapien
31	105	10.2	241	1 CDCA4_HUMAN	Q9BX18 homo sapien

32	104.5	10.1	740	2 Q9UPP2_HUMAN	Q9UPP2 homo sapien
33	101.5	9.8	728	1 P85B_HUMAN	Q00459 homo sapien
34	101.5	9.8	728	2 OSKAT5_HUMAN	Q05455 homo sapien
35	100.5	9.7	361	2 OS5TY1_CRYNE	Q55TY1 cryptococcu
36	100.5	9.7	361	2 OSKIF7_CRYNE	Q5KIF7 cryptococcu
37	100	9.7	846	2 Q6NVC9_BRARE	Q6NVC9 brachydanto
38	99	9.6	798	2 Q571B6_MOUSE	Q571B6 mus musculi
39	98.5	9.5	314	1 SOX12_MOUSE	Q04890 mus musculi
40	98	9.5	278	2 Q55XN0_CRYNE	Q55XN0 cryptococcu
41	98	9.5	517	2 Q9EEF8_NPYAG	Q9EEF8 anticarsia
42	98	9.5	562	2 Q8NAP0_HUMAN	Q8NAP0 homo sapien
43	98	9.5	768	2 Q5BEH7_EMENTI	Q5BEH7 aspergillus
44	98	9.5	1865	2 Q5VU37_HUMAN	Q5VU37 homo sapien
45	98	9.5	1884	2 Q4RWK6_TETNG	Q4RWK6 tetraodon n

ALIGNMENTS

RESULT 1

SRTD3_HUMAN STANDARD; PRT; 196 AA.

AC Q9UW9; Q9ECQ2; 01-FEB-2005 (Rel. 46, Created)
DT 01-FEB-2005 (Rel. 46, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE SRTD3 domain-containing protein 3 (Replication protein-binding trans-activator) (RPA-binding trans-activator).
GN Name=SRTD3; Synonyms=RB1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
OC Homo.

OK NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE, FUNCTION, SUBCELLULAR LOCATION, AND INTERACTION
RP WITH RPA2.
RX MEDLINE=20440390; PubMed=10982866; DOI=10.1093/nar/28.18.3478;
RA Cho J.M., Song D.J., Bergeron J., Benlamine N., Wold M.S.,
RA Alaoui-Jamali M.A.;
RT "RBT1, a novel transcriptional co-activator, binds the second subunit
of replication protein A";
RL Nucleic Acids Res. 28:3478-3485 (2000).
[2]

NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
TI TISSUE=Ovary, and Uterus; PubMed=12477932; DOI=10.1073/pnas.242603899;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Straussberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Hsieh F.,
RA Hopline R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loguettano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gamarale P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Wuzny D.W., Sodergren B.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Buterfield Y.S.N., Krzywicki M.T., Skalska U., Smallus D.B.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC -1- FUNCTION: Strong transcriptional co-activator.
CC -1- SUBUNIT: Interacts with RPA2.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: Contains 1 SRTD3 domain.

This Swiss-Prot entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.

CC -----
CC EMBL; AF192529; AAF05761.1; -; mRNA.
CC EMBL; BC014061; AAH14061.1; -; mRNA.
CC EMBL; BC050643; AAH50643.1; -; mRNA.
CC Ensembl; ENSG00000167565; Homo sapiens.
CC DR HGNC; HGNC:17931; SERTAD3.
CC DR InterPro; IPR009263; SERTA.
CC DR Pfam; PF06031; SERTA.1.
CC DR PROSITE; PS51053; SERTA.1.
CC DR Activator; Nuclear protein; Transcription; Transcription regulation.
CC FT DOMAIN 26 73 SERTA.
CC FT CONFLICT 2 2 V -> E (in Ref. 1).
CC SO SEQUENCE 196 AA; 21769 MW; 9C54APAL26F1AFL CRC64;
Query Match 99.3%; Score 1026; DB 1; Length 196;
Best Local Similarity 99.5%; Pred. No. 1.3e-70;
Matches 195; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MEGGKRGSDLEEEEREMWSPAGLOSYYQALLRISLDKVRSLGPAPSLRRHYLHN 60
DB 1 MVGGKRGSDLEEEEREMWSPAGLOSYYQALLRISLDKVRSLGPAPSLRRHYLHN 60
QY 61 TLQQLAALRLAPALPPPLFLGSDPFLSATIGSLIRELDTMDGTPEPPNVTPLG 120
DB 61 TLQQLAALRLAPALPPPLFLGSDPFLSATIGSLIRELDTMDGTPEPPNVTPLG 120
QY 121 LQNEVPPQDPVFLFALSSRYLGDGDLDFLIDITSVAKEKAPAPPEPPNLFCAQGS 180
DB 121 LQNEVPPQDPVFLFALSSRYLGDGDLDFLIDITSVAKEKAPAPPEPPNLFCAQGS 180
QY 181 WEMNELDHIMEITLIGS 196
DB 181 WEMNELDHIMEITLIGS 196

RESULT 2

SRTD3 MOUSE STANDARD; PRT; 197 AA.
AC Q9ERC3;
DT 01-FEB-2005 (Rel. 46, Created)
DT 01-FEB-2005 (Rel. 46, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE SERTA domain-containing protein 3 (Replication protein-binding trans-
DE activator) (RPA-binding trans-activator).
GN Name=Sertad3; Synonym=RBt1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
NCBI_TaxId=10090;
RN NUCLEOTIDE SEQUENCE.
RP Cho J.M., Alaoui-Jamali M.A.;
RT "Cloning of mouse RBt1 cDNA."
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
(2)

RN NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC STRAIN=C57BL/6J; TISSUE=Mammary gland;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold R.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Scheffer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonalant T.L., Casavant T.L., Scheetz T.B.,
RA Brownstein M.J., Ueda T.B., Toshlyuk S., Carninci P., Prange C.,
RA Rana S.S., Loggiano N.A., Peters G.J., Abramson R.D., Millady S.J.,
RA Boeck S.A., Mosman P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultky S.W.,

RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Farley J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalske U., Smalhus D.R.,
RA Schnerch A., Schein J.R., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- FUNCTION: Strong transcriptional co-activator (By similarity).
CC -1- SUBUNIT: Interacts with RPA2 (By similarity).
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -1- SIMILARITY: Contains 1 SERTA domain.

CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.

CC -----
CC EMBL; AF17202; AAC30951.1; -; mRNA.
CC EMBL; BC034866; AAH34866.1; -; mRNA.
CC Ensembl; ENSMUSG0000055200; Mus musculus.
CC DR MGI; MGI:2180697; Sertad3.
CC DR InterPro; IPR009263; SERTA.
CC DR Pfam; PF06031; SERTA.1.
CC DR PROSITE; PS51053; SERTA.1.
CC DR Activator; Nuclear protein; Transcription; Transcription regulation.
CC FT DOMAIN 27 74 SERTA.
CC SO SEQUENCE 197 AA; 21935 MW; CFF6FA4C0A91D7E7 CRC64;

Query Match 84.3%; Score 871; DB 1; Length 197;
Best Local Similarity 85.7%; Pred. No. 9e-59;
Matches 168; Conservative 8; Mismatches 18; Indels 2; Gaps 1;

QY 3 GGLKRRKSDLEEEEREMWSPAGLOSYYQALLRISLDKVRSLGPAPSLRRHYLHN 60
DB 2 GGLKRRKSDLEEEEREMWSPAGLOSYYQALLRISLDKVRSLGPAPSLRRHYLHN 61
QY 61 TLQQLAALRLAPALPPPLFLGSDPFLSATIGSLIRELDTMDGTPEPPNVTPLG 120
DB 62 TLQQLAALRLAPALPPPLFLGSDPFLSATIGSLIRELDTMDGTPEPPNVTPLG 121
QY 121 LQNEVPPQDPVFLFALSSRYLGDGDLDFLIDITSVAKEKAPAPPEPPNLFCAQGS 180
DB 122 PQNEIVSQADPVFLFALSSRYLGDGDLDFLIDITSVAKEKAPAPPEPPNLFCAQGS 181
QY 181 WEMNELDHIMEITLIGS 196
DB 182 WEMNELDHIMEITLIGS 197

RESULT 3

Q5BK27_RAT PRELIMINARY; PRT; 197 AA.
AC Q5BK27;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Hypothetical protein MGCL108974.
GN Name=MGCL108974;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Rattus.
NCBI_TaxId=10116;
RN NUCLEOTIDE SEQUENCE.

RP TISSUE=Liver;
RC TISSUE=Liver;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold R.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,

RA Altechul S.F., Zeeberg B., Buetow K.H., Scheifer C.F., Bhat N.K.,
 RA Hopline R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carrinci P., Prange C.,
 RA Raha S.S., Loguallano N.A., Peters G.J., Abramson R.D., Mullishy S.J.,
 RA Bosak S.A., McGowan P.J., McKernan K.U., Malek J.A., Gunaratne P.H.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren B.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.W.,
 RA Buterfield J.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.",
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN NUCLEOTIDE SEQUENCE.
 RP TISSUE=Liver;
 RC NIH MGC Project;
 RG Submitted (MAR-2005) to the EMBL/Genbank/DBJ databases.
 RL EMBL, BC091229; AAH91229.1; -, mRNA.
 DR InterPro: IPR011029; DRAFT_1like.
 DR InterPro: IPR003139; D retro matrix.
 DR InterPro: IPR000524; HTH GntR.
 DR InterPro: IPR009263; SERTA.
 DR Pfam: PF06031; SERTA; 1.
 DR PROSITE: PS51053; SERTA; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 197 AA; 21924 MW; 5F1586511550A9F2 CRC64;
 Query Match 83.1%; Score 858; DB 2; Length 197;
 Best Local Similarity 84.2%; Pred. No. 8,9e-58;
 Matches 165; Conservative 10; Mismatches 19; Indels 2; Gaps 1;
 Oy 3 GGLKKHSDLEBEE--ERWESPAGHSYQQAALNRISLDKVVORSIGPRAPSRRLVLLIN 60
 Db 2 GGLKKHSDLEBEEDEKDMSPYLRSTYQQAALLISLDKVVORSIGPRAPSRRLVLLIN 61
 Oy 61 TLQQAALRLAPALPEPEPLFLGEBDFSLATGSIIRLELDTSMDGTPEPPQNPVPLG 120
 Db 62 TLQQAALRLAPALPEPEPLFLGEBDFSLATGSIIRLELDTSMDGTPEPPQNPVPLG 121
 Oy 121 LQNEVPPPPPPYVFLALSSRYIGSDGLDFLIDITSANVKKPAPAPPPPPNLCARQS 180
 Db 122 PHDETMSQADPVFLALSSRYIGSDGLDFLIDITSANVKKPAPAPPPPPNLCARQS 181
 Oy 181 WEWNELDHIMEITIGS 196
 Db 182 WEWNELDHIMEITIGS 197
 RESULT 4
 SRTD1 MOUSE STANDARD; PRT; 236 AA.
 AC Q9JUI0; Q925B6; Q9D888;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE SBRP1 domain-containing protein 1 (Transcriptional regulator
 DE interacting with the PHD-bromodomain 1) (TRIP-B1) (CDK4-binding
 DE protein p34SE11) (SR-1).
 GN Name=Setd1; Synonym=Set1;
 OS Mus musculus (Mouse).
 OC Eukaryota; Euteleostomi; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Muridae; Murinae; Mus.
 NCBI_TaxId=10090;
 RN (1)
 RP NUCLEOTIDE SEQUENCE.
 RA Ohtani N., Hara E.;

RT "Cloning of mouse Set1-1 cDNA.",
 RL Submitted (DEC-1999) to the EMBL/Genbank/DBJ databases.
 RN NUCLEOTIDE SEQUENCE, FUNCTION, AND INTERACTIONS.
 RP MEDLINE=21231173; PubMed=11331592; DOI=10.1093/emboj/20.9.2273;
 RX Hsu S.-I., Yang C.M., Sim K.G., Hentchel D.M., O'Leary E.,
 RA Borventre J.V.,
 RT "TRIP-B1: a novel family of PHD zinc finger- and bromodomain-
 RT interacting proteins that regulate the transcriptional activity of
 RT E2F-1/DP-1.",
 RL EMBO J. 20:2273-2285(2001).
 RN NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
 RP STRAIN=C57BL/6J; Tissue=Embryo; and Small intestine;
 RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
 OKazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
 RA Nikaido I., Otsu N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
 RA Yagi K., Tomaru Y., Haegawa Y., Nogami A., Schonbach C., Gotohori T.,
 RA Baladrelli R., Hall D.P., Bult C., Hume D.A., Quackenbush J.,
 RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
 RA Blake J.A., Bradt D., Brusic V., Chochia C., Corbani L.E., Cousins S.,
 RA Dalla A., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
 RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
 RA Kongsava A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
 RA Maglott D.R., Marchionni L., McKenzie L., Miki H.,
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Pereira G.,
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
 RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
 RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Warshawsky Y., Wells C.,
 RA Wilming L.G., Wymshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carrinci P., Hayatsu N.,
 RA Hirozane-Kishikawa T., Kono H., Nakamura M., Sakazume N., Sato K.,
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Atakawa T., Fukuda S.,
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shingawa A.,
 RA Yasunishi A., Yoshino M., Waterson R., Lander E.S., Rogers J.,
 RA Birney E., Hayashizaki Y.,
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.",
 RL Nature 420:563-573(2002).
 RN NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
 RP MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnae.242603899;
 RX Klausner R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
 RA Altechul S.F., Zeeberg B., Buetow K.H., Scheifer C.F., Bhat N.K.,
 RA Hopline R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carrinci P., Prange C.,
 RA Raha S.S., Loguallano N.A., Peters G.J., Abramson R.D., Mullishy S.J.,
 RA Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren B.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Buterfield J.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.",
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -I- FUNCTION: Acts as E2F-responsive promoters to integrate signals
 CC provided by PHD- and/or bromodomain-containing transcription
 CC factors. Stimulates E2F-1/DP-1 transcriptional activity. Renders
 CC the activity of cyclin D1/CDK4 resistant to the inhibitory effects
 CC of p16(Ink4a).
 CC -I- SUBUNIT: Interacts with the PHD-bromodomain of TRIP1, TRIM28/TRIPB

CC and p300/CBP. Binds to DP1. Also interacts with CDK4.
 CC -1- SIMILARITY: Contains 1 SERTA domain.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 DR EMBL: AF218291; AAF27653.1; -; mRNA.
 DR EMBL: AF366400; AAK52829.1; -; mRNA.
 DR EMBL: AF366401; AAK52830.1; -; mRNA.
 DR EMBL: AK004022; BAB23130.1; -; mRNA.
 DR EMBL: AK008303; BAB25588.1; -; mRNA.
 DR EMBL: BC016077; AAI16077.1; -; mRNA.
 DR Ensemble: ENSMUSG0000008384; Mus musculus.
 DR MGI: MGI:1913438; Sertadi.
 DR InterPro: IPR009263; SERTA.
 DR Pfam: PF06031; SERTA; 1.
 DR PROSITE: PS51053; SERTA; 1.
 DR Transcription: Transcription regulation.
 DR DOMAIN: 38 SERTA.
 FT CONFLICT 9 10 KR -> NG (in Ref. 3; BAB25588).
 FT CONFLICT 12 12 E -> K (in Ref. 2; AAK52829).
 FT SEQUENCE 236 AA; 25136 MW; 989ADF829DE84C5 CRC64;
 SQ
 Query Match 21.0%; Score 217; DB 1; Length 236;
 Best Local Similarity 31.5%; Pred. No. 1.1e-08;
 Matches 76; Conservative 33; Mismatches 64; Indels 68; Gaps 12;
 QY 4 GLKRGSDLEEEERWE-----W-----SPAGLQS-----YQALLRISLDKVRSLGPR 48
 DB 5 GLKRR-----RESEETMEALSYDSCWLDPSHPAVAGTPPTVASSLFLDLSVYTLHSLQS 60
 QY 49 APSLRRLHLYNTLQQLQALRLAPALPREPL-----FLGEDSLSTISL 98
 DB 61 EPDRHLVLYVNTLRIRIQASME--PAPVLPPEPIQPPAPSVADSLASSDGLSMASTL 118
 QY 99 LREIDTSMGTEPPQNPVTPIGLQNEVPP-----QDPVFLRAL--SSRYLGDSDL 149
 DB 119 LEDLHIDLDLQAPQPA-----DEGPPGKSGISIRMLGLDLGLRATGLDGLDGLEG 172
 QY 150 FFLDIDTSAVERK-----PAR-----APRPPPHNLFCAPSGWENELDHWEILLG 195
 DB 173 LFEIDITSMYDSELWLPASEGLKPGEPNGPAKEBP-----ELDRAELDYLDLVAVG 224
 QY 196 S 196
 DB 225 T 225
 RESULT 5
 SRTD1 BRARE PRELIMINARY; PRT; 244 AA.
 ID 058B00 BRARE PRELIMINARY; PRT; 244 AA.
 AC 058B00;
 DT 10-MAY-2005 (TReMBLrel. 30, Created)
 DT 10-MAY-2005 (TReMBLrel. 30, Last sequence update)
 DT 10-MAY-2005 (TReMBLrel. 30, Last annotation update)
 DE Hypothetical LOC541508.
 GN Name=zgc:113340;
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 CC NCBI_TaxId=7955;
 CC [1]
 RN NUCLEOTIDE SEQUENCE.
 RP STRAIN=Singapore local strain; TISSUE=Embryo;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Struhsberg R.L., Felsing R.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schlier G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Uedeli T.B., Toshiyuki S., Carrinci P., Prange C.,
 RA Baka S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Roha S.A., McGowan P.J., Mckernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whitting M., Madan A., Young A.C., Snevchenko Y., Boulford G.G.,
 RA Blakeslee R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.B.,
 RA Schnerch A., Schein J.R., Jones S.J.M., Maira M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences".
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=Singapore local strain; TISSUE=Embryo;
 RG NIH MGC Project;
 RG Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC091808; AAI91808.1; -; mRNA.
 DR Hypothetical protein.
 SQ SEQUENCE 244 AA; 27015 MW; 8A02D0417D10D33A CRC64;
 Query Match 21.0%; Score 216.5; DB 2; Length 244;
 Best Local Similarity 28.7%; Pred. No. 1.3e-08;
 Matches 71; Conservative 36; Mismatches 73; Indels 67; Gaps 9;
 QY 4 GLKRR-HSDLEEEERWESPAGLQSYQALLRISLDKVRSLGPRAPSLRRLHLYNTL 62
 DB 5 GQKRLRRDDGAGASMS-----SQLQSYLDISMDKTKQDQALVERSLRSVLYNTL 57
 QY 63 QQLQALR-----LAPALPRE-----PLF-----LGED 88
 DB 58 RQVSEVVAQVDSLSPNLQSPKQDPPIQLSPVRNAPFLSLNAENVEDITGWTSSD 117
 QY 89 FSLSTIGSILREIDTSMGTEPPQNPVTPIGLQNEVPP-----QDPVFLRLSSRYLGD 144
 DB 118 FLSLSAISALIKDLDAIIDG--PSQR--SPIGSLNLPKCSGLRAKSYFRTAATCPGED 173
 QY 145 SGLDFFLDITSAVERK-----PAR-----APRPPPHNLFCAPSGWENELDHWEILLG 187
 DB 174 VSLDLMDIDSSVFEPEVNLLQGFSDLVKTLPSLSPSPASPSVSPGATWELQITE 233
 QY 188 HIMEILL 194
 DB 234 HMDPLM 240
 RESULT 6
 SRTD1 HUMAN STANDARD; PRT; 236 AA.
 ID SRTD1 HUMAN STANDARD; PRT; 236 AA.
 AC 09UHV2; 09UHV2;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 13-SEP-2005 (Rel. 48, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE SERTA domain-containing protein 1 (Transcriptional regulator
 DE interacting with the PHD-bromodomain 1) (TRIP-B1) (CDK4-binding
 DE protein p34SRT1) (SRT-1).
 GN Name=SRTAD1; Synonym=SR11;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominiidae;
 CC Homo.
 CC NCBI_TaxId=9606;
 CC [1]
 RN NUCLEOTIDE SEQUENCE, INTERACTION WITH CDK4, AND VARIANT ALA-31.
 RP MEDLINE=20047903; PubMed=10580009; DOI=10.1101/gad.13.22.3027;
 RX Sugimoto M., Nakamura T., Ohtani N., Hampson L., Hampson I.N.,
 RA Shinamoto A., Funichi Y., Okumura K., Niwa S., Taya Y., Hara E.,
 RT "Regulation of CDK4 activity by a novel CDK4-binding protein, p34(SRT-
 RT 1)".

RL Genes Dev. 13:3027-3033(1999).

RN [2]

RP NUCLEOTIDE SEQUENCE, FUNCTION, INTERACTIONS, AND VARIANT ALA-31.

RX MEDLINE:21231173; PubMed:11331592; DOI=10.1093/emboj/20.9.2273;

RA Hsu S.-I., Yang C.M., Sim K.G., Hentschel D.M., O'Leary E.,

RA Bonventre J.V.;

RT "TRIP-B": a novel family of PHD zinc finger- and bromodomain-

RT interacting proteins that regulate the transcriptional activity of

RT E2F-1/DP-1.";

RL EMBO J. 20:2273-2285(2001).

RN [3]

RP NUCLEOTIDE SEQUENCE (GENOMIC DNA).

RA Rieger M.J., Braun A.C., Montoya M.A., Chung M.-W., Nguyen C.P.,

RA Nguyen D.A., Livingston R.J., Poel C.L., Robertson P.D.,

RA Schackwitz W.S., Sherwood J.K., Witrak L.A., Nickerson D.A.;

RT "NIHNS-SNP", environmental genome project. NIHNS ES15478, Department

RT of Genome Sciences, Seattle, WA (URL: <http://esg.gs.washington.edu>).";

RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.

RN [4]

RP NUCLEOTIDE SEQUENCE (LARGE SCALE MRNA).

RP TISSUE=uterus;

RX MEDLINE:22389257; PubMed:12477932; DOI=10.1073/pnas.242603899;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Scheffen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Scheffer C.F., Heath N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heich F.,

RA Diatchenko L., Matsuda K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Ueda T.B., Toshiyuki S., Carninci P., Frange C.,

RA Raha S.S., Loggialano N.A., Peters G.J., Abramson R.D., Mullighy S.J.,

RA Bosak S.A., Mewen P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,

RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green B.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,

RA Buterfield V.S.N., Krzywinski M.J., Skalska U., Smailus D.B.,

RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human

RT and mouse cDNA sequences.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

CC -1- FUNCTION: Acts at E2F-responsive promoters to integrate signals

CC provided by PHD- and/or bromodomain-containing transcription

CC factors. Stimulates E2F-1/DP-1 transcriptional activity. Renders

CC the activity of cyclin D1/CDK4 resistant to the inhibitory effects

CC of p16(INK4a).

CC -1- SUBUNIT: Interacts with the PHD-bromodomain of TIF1, TRIM28/TIF1B

CC and p300/CBP. Binds to DPL. Also interacts with CDK4.

CC -1- SIMILARITY: Contains 1 SERPA domain.

CC -----

CC This Swiss-Prot entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use as long as its content is in no way modified and this statement is not

CC removed.

CC -----

CC DR EMBL, AF117959; AAF08349.1; -; mRNA.

CC DR EMBL, AF366402; AAK52831.1; -; mRNA.

CC DR EMBL, AY130860; AAM77800.1; -; Genomic_DNA.

CC DR EMBL, BC002670; AA02670.1; -; mRNA.

CC DR Ensembl, ENSG0000197019; Homo sapiens.

CC DR HGNc, HGNC:117932; SERPATD.

CC DR H-InvDB, HIX0015136; -;

CC DR GO, GO:0008284; P:positive regulation of cell proliferation; TAS.

CC DR GO, GO:0000079; P:regulation of cyclin dependent protein kina. .; TAS.

CC DR InterPro, IPR009263; SERPA.

CC DR Pfam, PF06031; SERPA. 1.

CC DR PROSITE, PS1053; SERPA. 1.

CC DR Polymorphism, Transcription; Transcription regulation.

CC DR KEGG, KEGG:015881; SERPA.

CC DR T -> A (in dbSNP:268687).

CC DR VARIANT 31 31 /FTid=VAR_015881.

SQ SEQUENCE 236 AA; 24704 MW; 8659505783BFE50B CRC64;

Query Match 20.6%; Score 213; DB 1; Length 236;

Best Local Similarity 32.6%; Pred. No. 2.3e-08;

Matches 76; Conservative 36; Mismatches 69; Indels 52; Gaps 13;

QY 4 GLKRRKSDLEEE---ERWESPAQLQSYQA-----LIRISLDKQSLGPRAP 51

DB 5 GLKRRKEEERKEPLAVDSW-WLDPGHTVAQAPPAVASSLPDLVLKHSLSQSEPD 63

QY 52 LRRHYLIHNTLOOLQALRLAPALPREPL-----FLGEEDFSLSATIGSLRE 101

DB 64 LRLHLVAVNTLRIQAS--MAPPAALPVPSPPAPSVADNLASSDALSHASWASLLE 121

QY 102 LPTSMQGTPEPPONPVYPLQIGNEVPP-----QPPVFLMAL--SSRYLGSGLDDPFL 152

DB 122 L-SHIEGLSQAPQ-----LADGPPGRIGGAAPSLGALDILGPRATGLDGLGLR 175

QY 153 DIDTSVAVERK---PARAPPEP-PHNLFCAPGSWE-----WNLIDHMEITLIGS 196

DB 176 DIDTSVYDNEIMAPASEGLKPGPED---QPGKEAPELDEAEDYLMVLT 225

RESULT 7

OS3GCO HUMAN

ID OS3GCO HUMAN PRELIMINARY; PRT; 236 AA.

AC OS3GCO;

DT 13-SEP-2005 (TREMBLrel. 31, Created)

DT 13-SEP-2005 (TREMBLrel. 31, Last sequence update)

DT 13-SEP-2005 (TREMBLrel. 31, Last annotation update)

DE SERPA domain containing 1 variant (Fragment).

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae;

OC Homo.

OX NCBI_Taxid=9606;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC TISSUE=Human small intestine;

RA Maruyama K., Sugano S.;

RT "Oligo-capping: a simple method to replace the cap structure of

RT eucaryotic mRNAs with oligobionucleotides.";

RL Gene 138:171-174(1994).

RN [2]

RP NUCLEOTIDE SEQUENCE.

RC TISSUE=Human small intestine;

RA Suzuki Y., Yoshitomo K., Maruyama K., Sugano S.;

RT "Construction and characterization of a full length-enriched and a 5'-

RT end-enriched cDNA library.";

RL Gene 200:149-156(1997).

RN [3]

RP NUCLEOTIDE SEQUENCE.

RC TISSUE=Human small intestine;

RA Suzuki Y., Sugano S., Totoki Y., Toyoda A., Takeda T., Sakaki Y.,

RA Tanaka A., Yokoyama S.;

RL Submitted (Apr-2005) to the EMBL/GenBank/DBJ databases.

DR EMBL, AK223011; BAD96731.1; -; mRNA.

DR EMBL, AK222970; BAD9690.1; -; mRNA.

DR NON TER 1

SQ SEQUENCE 236 AA; 24703 MW; 8659505783BFE50B CRC64;

Query Match 20.6%; Score 213; DB 2; Length 236;

Best Local Similarity 32.6%; Pred. No. 2.3e-08;

Matches 76; Conservative 36; Mismatches 69; Indels 52; Gaps 13;

QY 4 GLKRRKSDLEEE---ERWESPAQLQSYQA-----LIRISLDKQSLGPRAP 51

DB 5 GLKRRKEEERKEPLAVDSW-WLDPGHTVAQAPPAVASSLPDLVLKHSLSQSEPD 63

QY 52 LRRHYLIHNTLOOLQALRLAPALPREPL-----FLGEEDFSLSATIGSLRE 101

DB 64 LRLHLVAVNTLRIQAS--MAPPAALPVPSPPAPSVADNLASSDALSHASWASLLE 121

```
Qy 102 LDTSMDGTEPOMPVTPLGLONEVPP-----QPDVFLEAL--SSRYLGDSDGLDPFL 152
Db 122 L-SHIGKISQAPQP-----LADBGPGNSIGAAVPSLGLDLPATCLLDGGLDFE 175
Qy 153 DIDSNAVKE---PARAPPEP-PHNLFCAPGSWE-----NNELDHIMEITIGS 196
Db 176 DDTSMYDNEIMAPASBGLKPEPD---GPKKEBAPLEDEABLDYIMDLVGT 225

RESULT 8
06P771.RAT PRELIMINARY; PRT; 236 AA.
ID 06P771.RAT
AC 06P771.RAT
DT 05-JUN-2004 (TREMBlrel. 27, Created)
DT 05-JUN-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUN-2004 (TREMBlrel. 27, Last annotation update)
DE SERTA domain containing 1.
GN Name=Seratl;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Rattus.
OC NCBI_TaxID=10116;
OX [1]
RN NUCLEOTIDE SEQUENCE.
RP TISSUE=Prostate;
RC MEDLINE=23388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Straube R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins P.S., Wagner L., Shemen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buettow K.H., Scheffer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.B.,
RA Brownstein M.J., Ueda T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.O., Malek J.A., Gumaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren B.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.W.,
RA Butlerfield Y., Schin J.E., Jones S.J.M., Marra M.A.,
RA Scherch A., Schein J.R., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Prostate;
RA Director MGC Project;
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL, BC061808; AAH61808.1; mRNA.
DR InterPro, IPR009263; SERTA.
DR Pfam, PF06031; SERTA; 1.
SQ SEQUENCE 236 AA; 25419 MW; E8B7502E0473EDB8 CRC64;

Query Match 20.1%; Score 208; DB 2; Length 236;
Best Local Similarity 31.9%; Pred. NO. 5.4e-08;
Matches 76; Conservative 33; Mismatches 67; Indels 62; Gaps 13;

Qy 4 GUKRGHSDLEBBEREWE-----W-----SPAGIAGYQOALRLISLKVQSRIG 46
Db 5 GUKRR-----REBEETMESLSVDSWMLDGRPAVAGTPAVAS--SLPLDLSVKLHHSIR 58
Qy 47 PPAPELRHVTIHTTQOIALRLAPADLPPEPL-----FTGEEDPFSLSATIG 96
Db 59 QSEPLRHVIVVNTLRIOAGME--PTVLPPEPIOPTAPESADNFISSDAGLSASWA 116
Qy 97 SLRLDLDSMDGTEPOMPVTPLGLONEVPP-----QPDVFLEAL--SSRYLGDSDGL 147
Db 117 SLLEPLSHIDINQVPOQA-----DRGPPGRSVGCVLPNIGALDLPATGCLLDGGL 170
```

```
Qy 148 DDFELDITSAVKE---PARAPPEP-PHNLFCAPGSWE-----NNELDHIMEITIGS 196
Db 171 EGFLEDITSMYDSELMIPASBGLKPEPN---GPKKEBAPLEDEABLDYIMDLVGT 225

RESULT 9
09DC22.MOUSE PRELIMINARY; PRT; 244 AA.
ID 09DC22.MOUSE
AC 09DC22;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Mus musculus adult male kidney cDNA, RIKEN full-length enriched
DE library, clone:0610008D10 product:TRAP-BRI, full insert sequence.
GN Name=Seratl;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
OX [1]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=C57BL/6J; TISSUE=Kidney;
RC MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055590;
RA Kawai Y., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batilov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Glass C., King B., Kochiya H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schiml P.L., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarini R., Bash G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Giesbrecht S., Hill D., Hofmann U., Hume D.A., Kamuya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli U., Momberte P.,
RA Norone P., Ring B., Ringwald M., Rodriguez I., Sakamoto T.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-P.,
RA Suzuki H., Toyokawa K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Watanabe-Borja A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Kidney;
RA The PANOM Consortium;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Kidney;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes."
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Kidney;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
```

RA Kono H., Akiyama J., Nishi K., Katsunai T., Tashiro H., Itoh M.,
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 RA Fujiwara S., Inoue K., Togawa Y., Izawa Y., Ohara E., Matsubiki M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsunaga S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.,
 RT "RIKEN integrated sequence analysis (RISA) system-384-format
 RL sequencing pipeline with 384 multicapillary sequencer.";
 RA Genome Res. 10:1757-1771(2000).

NUCLEOTIDE SEQUENCE

RC STRAINE-C57BL/6J; T15SUB-Kidney;
 RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
 RA Arakawa T., Bono H., Caminici P., Fukuda S., Fukunishi Y., Furuno M.,
 RA Hasegaki T., Hara A., Hayatsu N., Hiramoto K., Hirooka T., Hori F.,
 RA Imocani K., Ishii Y., Itoh M., Izawa M., Kaakawa T., Kato H.,
 RA Kawai J., Kojima Y., Kono H., Kouda M., Koya S., Kurihara C.,
 RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
 RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
 RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
 RA Sogabe Y., Suzuki H., Tagami T., Tagawa A., Takehashi F., Tanaka T.,
 RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
 RA Mutamatsu M., Hayashizaki Y.,
 RA Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK002324; BAB2013.1; -; mRNA.
 DR Ensembl; ENSMUSG00000008384; Mus musculus.
 DR MGI; MGI:1913438; Sertadi.
 DR InterPro; IPR009263; SERTA.
 DR Pfam; PF06031; SERTA; 1.
 SQ SEQUENCE 244 AA; 26068 MW; 0F46F6419D362A6 CRC64;

Query Match 16.1%; Score 166; DB 2; Length 244;

Best Local Similarity 25.9%; Pred. No. 9.2e-05;
 Matches 67; Conservative 33; Mismatches 55; Indels 104; Gaps 13;

QY 4 GIKRKHSDLEEEERME-----W-----SPAGLS-----YQALLRISLDKVRQSLGPR 48
 DB 5 GIKRKN-----GEESEETWALSVDSWLDPSHPVAAGTPTVAASSLPDLISVVLTHSLKQS 60
 QY 49 APSLRRLVHNTLOQLAALRLAPALPPPL-----FLGEBDFSLATIGSI 98
 DB 61 EPDHLHLVLTVLTTRIIQASME--PAPVLPPEPIQPPAPSVADSLASDGLASMSL 118
 QY 99 LRELDTSMDGTPEPQ-----NPVTP-----LG-----LQNEV 125
 DB 119 LEDLNHIEDLNAPQADGPPGSGISGISPILGALDILGQPLAVCTMDRACLTYSI 178
 QY 126 PP-----OPDPVFLKALSSRYLGDSGLDFFLDIDTSAVEKEPARAPPEPNTLPCA 177
 DB 179 PPTTVTVVGYQP-----LRSVSAPEENGPAPK--EESP----- 208
 QY 178 PGSMWENELDHIMEITLGS 196
 DB 209 --ELDEAEITDLMIVLGV 225
 RESULT 10
 Q4T2W4_TESTNG PRELIMINARY; PRT; 244 AA.
 AC Q4T2W4;
 DT 13-SEP-2005 (TREMBlrel. 31, Created)
 DT 13-SEP-2005 (TREMBlrel. 31, Last sequence update)
 DE 13-SEP-2005 (TREMBlrel. 31, Last annotation update)
 DE Chromosome undetermined SCAFI0161, whole genome shotgun sequence.
 DE (Fragment).
 GN ORFNames=GSTENG00008134001;
 OS Tetradodon nigroviridis (Green puffer).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neocleosteii;
 OC Acanthomorphi; Acanthopterygii; Percomorphi; Tetraodontiformes;
 OC Tetraodontidae; Tetraodontidae; Tetraodon.
 OK NCB1_Taxid=99883;
 [1]

RP NUCLEOTIDE SEQUENCE.
 RA Jallion O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
 RA Maucell E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
 RA Nicud S., Jaffe D., Fisher S., Lutfalla G., Dosset C., Segurens B.,
 RA Dastiva C., Salanoubat M., Levy M., Boudet N., Cascellano S.,
 RA Anthouard V., Dubin C., Castelli V., Katinka N., Vacherie B.,
 RA Bismont C., Skalli Z., Catolico L., Poulain J., De Berardinis V.,
 RA Parra G., Lardier G., Chapple C., McKernan K.J., MEBwan P., Bosak S.,
 RA Kellis M., Volff J.N., Guiso R., Zody M.C., Mestrov J.,
 RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
 RA Laudet V., Schachter V., Quetier F., Sautin W., Scarpelli C.,
 RA Winkler P., Lander E.S., Weissbach J., Roest Croollis H.,
 RT "Genome duplication in the teleost fish Tetradodon nigroviridis reveals
 RL the early vertebrate proto-karyotype.";
 RA Nature 431:946-957(2004).

NUCLEOTIDE SEQUENCE

RP NUCLEOTIDE SEQUENCE.
 RG Genoscope; Whitehead Institute Centre for Genome Research;
 RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
 CC -!- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL; CAE010161; CAP92768.1; -; genomic DNA.
 FT NON TER 1
 FT 1
 SQ SEQUENCE 244 AA; 26476 MW; 796E0C82A3644A05 CRC64;

Query Match 15.8%; Score 163.5; DB 2; Length 244;

Best Local Similarity 29.8%; Pred. No. 0.00014;
 Matches 68; Conservative 27; Mismatches 86; Indels 47; Gaps 8;

QY 4 GIKRKHSDLEEE-----EERWESWSPAGL-----QSYQALLRISLDKVRQSLGPRASL 52
 DB 29 GYKRWKSLDEBPAALPAADKXSE--PDGFLGPPSSVWLLGLCKLQHQTVGVSL 86
 QY 53 RRVVLHNTLOQLAALRLAPALPPPLFLGEBDFSLATIGSILRELDTSMDGTPEP 112
 DB 87 RRSVLLINTLQIQEISQDGMGTCAPILGAVGTQDSC-----PLRQDLPVTCGRGQPA 141
 QY 113 QNPVPLGLQNEVPPQPPVFLKALS--RYLGDSGLDFFLDIDTSAVEKE----- 162
 DB 142 RR-----SSKRTRQQAATLFGDANANAVGYSIDLDFIDIDTSMWESSDLPASWAVG 196
 QY 163 ---PARAPPEPNTLFCAPGS-----WENELDHIMEITLGS 196
 DB 197 SLWPVSISLMADEDLKVCSSHAKSGLSQSCMDLNEIDQIMEITLVS 244

RESULT 11

Q4S6F1_TESTNG PRELIMINARY; PRT; 275 AA.
 AC Q4S6F1;
 DT 13-SEP-2005 (TREMBlrel. 31, Created)
 DT 13-SEP-2005 (TREMBlrel. 31, Last sequence update)
 DT 13-SEP-2005 (TREMBlrel. 31, Last annotation update)
 DE Chromosome 10 SCAFI4728, whole genome shotgun sequence.
 DE (Fragment).
 GN ORFNames=GSTENG00023321001;
 OS Tetradodon nigroviridis (Green puffer).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neocleosteii;
 OC Acanthomorphi; Acanthopterygii; Percomorphi; Tetraodontiformes;
 OC Tetraodontidae; Tetraodontidae; Tetraodon.
 OK NCB1_Taxid=99883;
 [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Jallion O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
 RA Maucell E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
 RA Nicud S., Jaffe D., Fisher S., Lutfalla G., Dosset C., Segurens B.,
 RA Dastiva C., Salanoubat M., Levy M., Boudet N., Cascellano S.,
 RA Anthouard V., Dubin C., Castelli V., Katinka N., Vacherie B.,
 RA Bismont C., Skalli Z., Catolico L., Poulain J., De Berardinis V.,
 RA Parra G., Lardier G., Chapple C., McKernan K.J., MEBwan P., Bosak S.,
 RA Kellis M., Volff J.N., Guiso R., Zody M.C., Mestrov J.,
 RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
 RA Laudet V., Schachter V., Quetier F., Sautin W., Scarpelli C.,
 RA Winkler P., Lander E.S., Weissbach J., Roest Croollis H.,
 RT "Genome duplication in the teleost fish Tetradodon nigroviridis reveals
 RL the early vertebrate proto-karyotype.";
 RA Nature 431:946-957(2004).

RA Parra G., Lartier G., Chapelle C., McKernan K.J., McEwan P., Bosak S.,
 RA Kellie M., Volff J.N., Guigo R., Zody M.C., Mesirov J.,
 RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
 RA Lander V., Schachter V., Queller F., Saurin W., Scarpelli C.,
 RA Winkler P., Lander E.S., Weissbach J., Roest Croliis H.,
 RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
 RT the early vertebrate proto-karyotype.";
 RL Nature 431:946-957 (2004).
 RN
 RN NUCLEOTIDE SEQUENCE.
 RP Genoscope; Whitehead Institute Centre for Genome Research;
 RG Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
 CC -1- CATTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 CC EMBL, CAAB01014728; CAG03781.1; -; Genomic_DNA.
 DR NON TER 275 275
 FT SEQUENCE 275 AA; 29572 MW; 9CF6A854C3C1349 CRC64;
 SQ
 Query Match 14.0%; Score 145; DB 2; Length 275;
 Best Local Similarity 24.9%; Pred. No. 0.0043;
 Matches 72; Conservative 35; Mismatches 66; Indels 116; Gaps 12;
 QY 4 CLKRGHSLDEERENWEMSPAGLO-----SY---QALLRISLDKQKRSKL 45
 DB 5 GVKRFPSDAGEBA-----APSGEGPPATSAARTLSSSYSLQKSLDMSLITLQCH 60
 QY 46 GPRAPSLRRVLIHNTLOOLQ-----AALRLAPALPPEPLFGEEDFSL 92
 DB 61 MLVENVLCRSVLIANTVROIOEBMTQDQWQIMTALAAQCPA-----DRIVA 109
 QY 93 ATIGSLRLDLTSMGTEPPONPVPLGIQ-----NEVPPOD-----PVFLPA 136
 DB 110 SEV--LCRGTDAAPAGQSPK-PFSVYGLBEGYHSEVVMEDMEPEVTMTSLSPVSPQL 166
 QY 137 LSSRLG-----DSGLDFDLDTITSAVKEP 163
 DB 167 SSASVYLGQPGMGPCWEEBEDVTGQVFTFKHPAPPDPALBELTSDVDPSYDDT 226
 QY 164 A-----RAPPEPPHNLFCAPGSW-----ENNELDHIMEITL 194
 DB 227 VLTGMSAPKMGPPYDLLESLSHGPTALLSSSQCSDLNELDHIMEITL 275
 RESULT 12
 OSRIK3 BRARE PRELIMINARY; *PRT; 229 AA.
 ID OSRIK3 BRARE PRELIMINARY; *PRT; 229 AA.
 AC OSRIK3;
 DT 01-FEB-2005 (TREMBlrel. 29, Created)
 DT 01-FEB-2005 (TREMBlrel. 29, Last sequence update)
 DT 01-FEB-2005 (TREMBlrel. 29, Last annotation update)
 DE Novel protein (Fragment).
 GN Name=sl:dkey-177p2.6; Synonyms=OTDAR0000005081;
 GN ORFNames=DKEY-177P2.6-001;
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OK NCBI_TaxId=7955;
 RN
 RN NUCLEOTIDE SEQUENCE.
 RP Loveall J.;
 RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL, BX248515; CAI20741.1; -; Genomic DNA.
 DR ZFIN, ZDB-GENE-030131-5526; sl:dkey-177p2.6.
 DR InterPro, IPR009263; SERPA.
 DR Pfam, PF06031; SERPA.
 FT NON TER 229 229
 FT SEQUENCE 229 AA; 25097 MW; F7936F752874AA1F CRC64;
 SQ
 Query Match 14.0%; Score 144.5; DB 2; Length 229;
 Best Local Similarity 29.2%; Pred. No. 0.0038;
 Matches 57; Conservative 14; Mismatches 61; Indels 63; Gaps 6;

QY 4 CLKRGHSLDEERENWEMSPAGLOSY---QALLRISLDKQKRSPPASLRVLIHN 60
 DB 57 GVKRLDSTCEB-----PAQDLPYPOQRQVLVDLCIDKIQSCORRAEPILHRVLLAN 108
 QY 61 TLQOLQALR-----LAPA-----PALPEPL----- 82
 DB 109 TLKQIQEMRQSGTCLPPAVLQPSLPITQPRHVPFELPPVFLDOPALVTALSPFLM 168
 QY 83 FLGERDFSLSATIGSLRELDTSMDGTEPPONPVPLGIQNEVPPODPFLEALSRYL 142
 DB 169 TADEERELGGTETEMTLLPLSLAGEDNTKSDLLFGSFRTN-----STSYL 214
 QY 143 GDSGLDFFLIDITS 157
 DB 215 TDLALDDIFEDIDTS 229
 RESULT 13
 SRTD2 MOOSE STANDARD; PRT; 309 AA.
 ID SRTD2 MOOSE STANDARD; PRT; 309 AA.
 AC Q9JUG5; Q8C609; Q9JWL3; Q925B5;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE SERPA domain-containing protein 2 (transcriptional regulator
 DE interacting with the PHD-bromodomain 2) (TRIP-Br2).
 GN Name=Seratad2; Synonyms=K1aa0127; ORFNames=MNCB-1504;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muricidae; Muridae; Murinae; Mus.
 OK NCBI_TaxId=10090;
 RN
 RN NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
 RC STRAIN=C57BL/6; TISSUE=Brain;
 RA Osada N., Kusuda J., Tanuma R., Ito A., Hirata M., Sugano S.,
 RA Hashimoto K.;
 RT "Isolation of full-length cDNA clones from mouse brain cDNA library
 RT made by oligo-capping method.";
 RL Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.
 RN
 RN NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
 RP STRAIN=C57BL/6J; TISSUE=Testis;
 RC MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
 RA Nishida I., Oeato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gotohori T.,
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
 RA Schiml L.M., Kanapin A., Matuda H., Batilov S., Beisel K.W.,
 RA Blake J.A., Bradt D., Brusic V., Chochia C., Corbani L.E., Cousins S.,
 RA Dalla B., Dragan T.A., Fletcher C.F., Forester A., Fraser K.S.,
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
 RA Grimmond S., Guerinich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 RA Kanai A., Kawaji H., Kawasawa Y., Kedzierzki R.M., King B.L.,
 RA Kongawa A., Kurochkin I.V., Lee Y., Lehman B., Lyons P.A.,
 RA Maglott D.R., Maltaise L., Marchionni L., McKenzie L., Miki H.,
 RA Nagashima T., Numata K., Okido T., Pavan M.J., Petosa G., Pesole G.,
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramchandran S.,
 RA Ravasi T., Reed J.C., Reed D.O., Reid J., Ring B.Z., Ringwald M.,
 RA Sandelin A., Schneider C., Sempile C.A., Setou M., Shimada K.,
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
 RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
 RA Wilming L.G., Wyrshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
 RA Hirozane-Kishikawa T., Kono H., Nakamura M., Sakazume N., Sato K.,
 RA Shitaki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 RA Birney E., Hayashizaki Y.;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";

RL Nature 420:563-573 (2002).
[3]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RP TISSUE=Kidney;
RC MEDLINE=2238657; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Struhsberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shemmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.O., Ueda T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullighan S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Huijck S.W.,
RA Villalón D.K., Muzny D.M., Sodergren B.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green B.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Buterfield Y.S.N., Krzywiński M.J., Skalska U., Smallus D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
[4]
RP NUCLEOTIDE SEQUENCE OF 129-309.
RX MEDLINE=21231173; PubMed=1331592; DOI=10.1093/emboj/20.9.2273;
RA Hsu S.-I., Yang C.M., Sim K.G., Hentchel D.M., O'Leary E.,
RA Bonventre J.V.;
RT "TRP-B1: a novel family of PHD zinc finger- and bromodomain-
RT interacting proteins that regulate the transcriptional activity of
RT E2F-1/DP-1.";
RL EMBO J. 20:2273-2285 (2001).
CC -!- FUNCTION: Acts as E2F-responsive promoters to integrate signals
CC provided by PHD- and/or bromodomain-containing transcription
CC factors (by similarity).
CC -!- SIMILARITY: Contains 1 SERTA domain.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL, AB041541; BAA95026.1; -; mRNA.
CC EMBL, AK076787; BAC36480.1; -; mRNA.
CC EMBL, BC014726; AAH14726.1; -; mRNA.
CC EMBL, AF366403; AAK52832.1; -; mRNA.
CC EMBL, AF366403; AAK52832.1; -; mRNA.
CC Ensembl, ENSMUSG0000049800; Mus musculus.
CC MGI, MGI:1931026; Sertad2.
CC InterPro, IPR009263; SERTA.
CC Pfam, PF06031; SERTA.1.
CC PROSITE, PS51053; SERTA.1.
CC Transcription; Transcription regulation.
CC -----
CC DOMAIN 33 80
CC FT CONFLICT 138 140 CTL -> FTF (in Ref. 4).
CC FT CONFLICT 142 142 A -> S (in Ref. 3).
CC FT CONFLICT 145 145 P -> S (in Ref. 4).
CC FT CONFLICT 234 234 T -> A (in Ref. 1).
CC SQ SEQUENCE 309 AA; 33312 MW; D4178688FDDF8F00 CRC64;
Query Match 13.6%; Score 140.5; DB 1; Length 309;
Best Local Similarity 24.8%; Pred. No. 0.011;
Matches 77; Conservative 29; Mismatches 85; Indels 119; Gaps 13;

QY 87 ----- EDFSLSATIGSILRELDTSMDG 108
DB 121 ACTTPASLLEDDNDPCTLQAVHPAPTRLSAALPAEKDSFSSALDEI -BELCPTSTS 179
QY 109 TEPPO--PVTPPLGQNEVPP-----PDVPLEAL-----SSRYLDSDG 148
DB 180 TEAHTAABEGKGTSSSSSVQKPGREGRTDSRFMDLSLGNPEITTTSTGFLTDLTLD 239
QY 149 D-FPLDIDTSAVEKEPA-----RAPPE-----PPHNLFCAPG---SWEWNE 186
DB 240 DILFADIDTSTWDPPTCSASGTASKMAPVSADLTKTLAPVSNQPVAPSGPFKMDLT 299
QY 187 DHIMEIILGS 196
DB 300 DHIMEIVLGS 309
RESULT 14
QSSSES MOUSE PRELIMINARY; PRT; 309 AA.
ID QSSSES5
AC QSSSES5;
DT 10-MAY-2005 (TREMBlrel. 30, Created)
DT 10-MAY-2005 (TREMBlrel. 30, Last sequence update)
DE Novel protein similar to cell division cycle associated 4 Cdc44.
GN Name=RP23-452C23.1; ORFNames=RP23-452C23.1-001;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OC NCBI_Taxid=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Beasley H.;
RL Submitted (FEB-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL, AL663115; CAI5740.1; -; Genomic DNA.
KW cell division.
SQ SEQUENCE 309 AA; 33312 MW; D4178688FDDF8F00 CRC64;
Query Match 13.6%; Score 140.5; DB 2; Length 309;
Best Local Similarity 24.8%; Pred. No. 0.011;
Matches 77; Conservative 29; Mismatches 85; Indels 119; Gaps 13;

QY 2 EGGLEKSHSDLEBERWEMSPAGLOS-----YQALRLSLDKVQSLGRAPSLRH 55
DB 4 KGG-KRKEDEHEDGEGKIVSPDGSRYVTLQRTIFNLSLMLKYNHRPLTEPSLQKT 62
QY 56 VLIHNTLOQLQALRL-----APAPLPPEPLFLG----- 86
DB 63 VLIHNTLRRIQELKQEGSLRAPATPSPGSPNSLSDSYQEAAPP--PHPCDLGSGTTPLE 120
QY 87 ----- EDFSLSATIGSILRELDTSMDG 108
DB 121 ACTTPASLLEDDNDPCTLQAVHPAPTRLSAALPAEKDSFSSALDEI -BELCPTSTS 179
QY 109 TEPPO--PVTPPLGQNEVPP-----PDVPLEAL-----SSRYLDSDG 148
DB 180 TEAHTAABEGKGTSSSSSVQKPGREGRTDSRFMDLSLGNPEITTTSTGFLTDLTLD 239
QY 149 D-FPLDIDTSAVEKEPA-----RAPPE-----PPHNLFCAPG---SWEWNE 186
DB 240 DILFADIDTSTWDPPTCSASGTASKMAPVSADLTKTLAPVSNQPVAPSGPFKMDLT 299
QY 187 DHIMEIILGS 196
DB 300 DHIMEIVLGS 309
RESULT 15
QSGMB1_XENLA
ID QSGMB1_XENLA PRELIMINARY; PRT; 248 AA.
AC QSGMB1;

Search completed: February 6, 2006, 10:01:09
 Job time : 257 secs

```

DT 05-JUL-2004 (TRENBLrel. 27, Created)
DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
DE MGCS2098 protein.
GN Name=MGCS2098;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OC NCBI_TaxID=8355;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Kidney;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RX Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RX Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RX Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RX Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RX Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Schetz T.E.,
RX Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RX Rah S.S., Loguellano N.A., Peters G.J., Abramson R.D., Mullahy S.D.,
RX Boeck S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RX Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RX Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RX Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RX Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RX Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RX Butlerfield Y.S.N., Krzywinski M.I., Skalski U., Smalls D.B.,
RX Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Kidney;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RX Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RX Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT Initiative."
RT Dev. Dyn. 225:384-391(2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Kidney;
RX Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
RX EMBL; BC074195; AAH74195.1; -; mRNA.
RX InterPro; IPR009263; SERTA.
RX Pfam; PF06031; SERTA; 1.
SQ SEQUENCE 248 AA; 27319 MW; E703D7E9105DF08C CRC64;

Query Match 13.2%; Score 136.5; DB 2; Length 248;
Best Local Similarity 30.9%; Pred. No. 0.017;
Matches 58; Conservative 25; Mismatches 58; Indels 47; Gaps 8;

QY 17 ERMWSPAGLQSYQQAALISLKVORSLGPAPSLRRHYLHNTLQQLQALRL--AP 73
DB 10 ERECCSIPAIQSH--CLMNISLVKLRSLRHVEPDLRHFLVAVNTIKRLQGNLQVEQCNP 67
QY 74 -----APALPEPPLFGLGBDPSLSATIGSILRELDTSMDGTBPPO 113
DB 68 DMMKTSBECTRSALVVPBSKQPALNTEBDPLSSMDASTLSYSTILEDLN-NPBGLS--S 124
QY 114 NPVTPLGLQNEVPPQPDVFLER-----LSRY-----LQPSGLDDFFLDID 155
DB 125 SPLPQIEDDQLCAFKANPVYSGAEDMVYKLASSSSLSLSSSPYLLGKNTGDN-LEDIFFEDID 183
QY 156 TSAVERKEP 163
DB 184 TSMYSDP 191

```



```

; GENERAL INFORMATION:
; APPLICANT: JURCEIC, ROLAND
; TITLE OF INVENTION: HEP, A NOVEL GENE WITH A ROLE IN HEMATOPOIETIC AND NEURAL DEVELO
; FILE REFERENCE: 39532-176599
; CURRENT FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: US 60/268,923
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patent version 3.0
; SEQ ID NO 4
; LENGTH: 241
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-076-069-4

Query Match          10.2%; Score 105; DB 2; Length 241;
Best Local Similarity 25.9%; Pred. No. 0.032;
Matches 65; Conservative 31; Mismatches 79; Indels 76; Gaps 13;

QY 4 GUKRGSDLEBEEBEMWSPAGI-Q---SY---QALLRISLDKVRSLGPRAPSLRRHYL 57
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 5 GUKKRCVGHEDV---EGALAGLKVSSYSLQROSLDMSLVKQLCHMLVPPNLCRSVL 61
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 58 IHNTLQOLAAL-----RLAPAPALPPEPLFLGEBDFSLATIGSLIRELDTSMGTEP 111
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 62 IANTVROQEBMTQDGTWRTVAPOAERAPL-----DRLVST---ELLCRAAMGQEGANP 113
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 112 -----PQNPVT-----PLGLONEVPPQDP-----VLELSSRY 141
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 114 APGLDGHGTGCPVSDLCPTVSAQAPRHQLQSSAMENDPKEKNGSPHKSIDQIFETLETK- 172
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 142 LGDSGLDPEFLDIDTSAAVEKP-----AR-----APPEPHNLCAPGSMW 183
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 173 -NPSCHBELFSVDSPYDLDIVLVGMGAGAPGCGEGLAGAPATPGSSSCXG---DL 228
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 184 NELDHIMEITL 194
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 229 GELDHVEITLV 239

RESULT 3
US-10-104-047-3214
; Sequence 3214, Application US/10104047
; Patent No. 6943241
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. 6943241el full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: Patent ver. 2.1
; SEQ ID NO 3214
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-3214

Query Match          9.5%; Score 98; DB 2; Length 562;
Best Local Similarity 32.8%; Pred. No. 0.45;
Matches 40; Conservative 7; Mismatches 39; Indels 36; Gaps 6;

QY 48 RAPSRLRHYLIHN-----TIQOLAALRLAPAPALPPEPLFLGEBDFSLATIGSLIR 100
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 452 RAYSLLRHQCRABLERAAALQALQAPTSPP--PPPLKABQEBEGLPLPLANI-- 507
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 101 ELDTSMGTEPPQNPVTYGLONEVPPQDPVFLLELSSRYLGDSGLDPEFLDIDTSAAVE 160
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 508 -----KEBPPSPGTP-----PQSPAP-PVFLSA-----SCFQSDHSAPV 542
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```

QY 161 KE 162
DB 543 ME 544

RESULT 4
US-09-949-016-8918
; Sequence 8918, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8918
; LENGTH: 529
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-8918

Query Match          9.2%; Score 95.5; DB 2; Length 529;
Best Local Similarity 23.4%; Pred. No. 0.72;
Matches 46; Conservative 25; Mismatches 71; Indels 55; Gaps 9;

QY 20 EWSPA-----GLOSQOAL-----LRISLDKVRSLGPRAPSLRRHYLIH 59
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 174 QMDTALADGIXFLLAPALPALTVPASAKARALREAAAGPAPLPPPTLPLHATLIR 233
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 60 NTLQOLAALRLAPAPALPPEPLFLGEBDFSLATIGSLIRELDTSMGTEPPQNPVTPL 119
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 234 FLQHLGRVVASRAPA-----LGPVVALGATFGPL-----LRAPPPSP-PG 277
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 120 GLONEVPPQDP--PVFLLELSSRYLGDSGLDPEFLDIDTSAAVEKPAPAPPP----- 170
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 278 GAPDSESPDPFPALVETLQEHLEBQ-----EVAPPALPPPKAKPAPVLANG 330
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 171 -PHNLCAPGSMWENEL 186
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 331 SPPSL--QDAEWYMDI 345

RESULT 5
US-08-259-451-3
; Sequence 3, Application US/08259451
; Patent No. 6406841
; GENERAL INFORMATION:
; APPLICANT: Lee, Helen H.
; APPLICANT: Swanson, Patricia A.
; APPLICANT: Idler, Kenneth B.
; APPLICANT: Rosenblatt, Joseph D.
; APPLICANT: Chen, David W.
; APPLICANT: Golde, David W.
; APPLICANT: Robertson, Eugene F.
; APPLICANT: Stephens, John B.
; APPLICANT: Chan, Emerson W.
; APPLICANT: Buytendorp, Mark H.
; APPLICANT: Johnson, Joan E.
; APPLICANT: Motley, Cheryl T.
; APPLICANT: Peterson, Bryan
; APPLICANT: Edwards, Michelle
; APPLICANT: Guidinger, Peggy
; APPLICANT: Tate, Cynthia
```



```

Qy 47 PRAPSLRRHVLINHTLOQLALRLAPAP-----LPPPLFGEDPFSLSA 93
Db 13 PKAP-----RGSTHWMINFLQAAAYRLQPRPSDFQOLRLFKLAKTPIWMLPIDYSLA 69
Qy 94 TI-----GSLRELDSMDGTEPPONVTPLG---LQNEVPPODPVFLA-LSSRYLG 143
Db 70 SLIRKYPGRVVEITINIVKQVBSADAPVPTPICTTTPPPPPSPAAHPPPY-- 127
Qy 144 DSGLDLFDLIDTSAVEKEPARAPPEPHNLFCAPGSWENMLDHIMEIIGS 196
Db 128 -----VEPTTQCFPIILHPPGAP-----SARHPQMDLQAIKQEVSS 166

```

RESULT 8

```

US-09-413-814-28
Sequence 28, Application US/09413814
Patent No. 6225064

```

```

GENERAL INFORMATION:
APPLICANT: Gesellschaft fuer Biotechnologische Forschung mbH
APPLICANT: Bristol-Myers Squibb, Co.
APPLICANT: Beyer, Stefan
APPLICANT: Bloecker, Helmut
APPLICANT: Brandt, Petra
APPLICANT: Cino, Paul M
APPLICANT: Dougherty, Brian A
APPLICANT: Goldberg, Steven L
APPLICANT: Hofle, Gerhard
APPLICANT: Mueller, Joachim
APPLICANT: Reichenbach, Hans
TITLE OF INVENTION: DNA sequences for enzymatic synthesis of polyketide or
FILE REFERENCE: PCT/US 99/23535
CURRENT FILING DATE: 1999-10-07
EARLIER APPLICATION NUMBER: US/09/413,814
EARLIER FILING DATE: 1998-10-09
NUMBER OF SEQ ID NOS: 107
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 28
LENGTH: 2618
TYPE: PRT
ORGANISM: Sorangium cellulosum
US-09-413-814-28

```

```

Query Match 9.0%; Score 93; DB 2; Length 2618;
Best Local Similarity 25.5%; Pred. No. 10;
Matches 55; Conservative 18; Mismatches 79; Indels 64; Gaps 11;

```

```

Qy 31 QALLR--ISLD-----KVQSLGPPAP--SLRRVLIHNTLQQLQA 67
Db 974 QALLREAIGLDEPFQAGNGSFLIRLKLKLSAFKSPITDLEFHTSIRSQAEMLSG 1033
Qy 68 ALRLAP-APALPPEPLFGEDPFSLSATIGSLRELDSMDGTEPPONVTPVLGQN 123
Db 1034 SSVETPLGAVYQPPAAAQVASSAKSPGERGAATSSGLTAOPPHFRIVAYIGLAG 1093
Qy 124 EYPPQPD-PVFLBAL-----SSRYLGSGLD-----DFPLD 153
Db 1094 RPPAAPDDDAFELITBGRGIRFSGQAEIRBEGIDANKRIACHYVPAKGFIDRADHF-D 1152
Qy 154 IDTSAVEKEPAR-APPEPHNLFCAPGSWENMLDH 188
Db 1153 ADFGCIPPDAETIDPQIRLLLECC-----VNALEH 1183

```

RESULT 9

```

US-10-076-069-6
Sequence 6, Application US/10076069
Patent No. 6872812

```

```

GENERAL INFORMATION:
APPLICANT: JURECIC, ROLAND
APPLICANT: NACHTMAN, RONALD
TITLE OF INVENTION: HBEP, A NOVEL GENE WITH A ROLE IN HEMATOPOIETIC AND NEURAL DEVELOPMENT

```

```

FILE REFERENCE: 39532-176599
CURRENT APPLICATION NUMBER: US/10/076,069
CURRENT FILING DATE: 2002-02-15
PRIOR APPLICATION NUMBER: US 60/268,923
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 11
SOFTWARE: Patentin version 3.0
SEQ ID NO 6
LENGTH: 113
TYPE: PRT
ORGANISM: Mus musculus
US-10-076-069-6

```

```

Query Match 9.0%; Score 92.5; DB 2; Length 113;
Best Local Similarity 45.3%; Pred. No. 0.19;
Matches 29; Conservative 11; Mismatches 21; Indels 3; Gaps 2;

```

```

Qy 4 GLKRKSHLEEREERWSPR-GLQSYQALLRISLDVQKSLGPRAPSLRRHVLINHTL 62
Db 5 GLKRKYQDSRGVSGFVPSYIQ--RQSLDMSLVQLQCHMLVEPNLCRSVLIANTV 62
Qy 63 QQLQ 66
Db 63 RQIQ 66

```

RESULT 10

```

PCT-US95-04910-13
Sequence 13, Application PC/TUS9504910

```

```

GENERAL INFORMATION:
APPLICANT: The Government of the United States of America as represented by the Secretary, Department of Health and Human Services
TITLE OF INVENTION: ISOLATION AND CHARACTERIZATION OF A NOVEL PRIMATE T-CELL LYMPHOTROPIC VIRUS AND THE USE OF THIS VIRUS
TITLE OF INVENTION: OR COMPONENTS THEREOF IN DIAGNOSTIC ASSAYS
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESS: MORGAN & FINNEGAN
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154

```

COMPUTER READABLE FORM:

```

MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/04910
FILING DATE: 21-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US08/231,526
FILING DATE: 22-APR-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: WILLIAM S. FEILER
REGISTRATION NUMBER: 26,728
REFERENCE/DOCKET NUMBER: 2026-4125PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 432 amino acids
TYPE: amino acids
STRANDEDNESS: unknown
TOPOLOGY: unknown

```


PRIOR FILING DATE: 1999-06-10
PRIOR APPLICATION NUMBER: US 60/097,334
PRIOR FILING DATE: 1998-08-18
NUMBER OF SEQ ID NOS: 72
SOFTWARE: Patentin version 3.0
SEQ ID NO 16
LENGTH: 370
TYPE: PRT
ORGANISM: Rattus norvegicus
US-10-192-361-16

Query Match 8.8%; Score 90.5; DB 2; Length 370;
Best Local Similarity 23.5%; Pred. No. 1.4;
Matches 43; Conservative 29; Mismatches 62; Indels 49; Gaps 8;

QY 33 LRLSLDLY-GRSLGPRAPSLRRHVLINHTLOQALRLAPALP-----EPLFLG 85
DB 102 MNWHLKVARREIGTLA-----TVRLPPSKVLPPELPLTYCKRFL--- 147
QY 86 EEDSLATISILRELDTSMDGT-----EPPONFVPLGLONEVPO-PDPVLEAL- 137
DB 148 --NFACLDVGHGVADLSTQSLRTGTLRSKSIKAPATASATIGRPRIPEVQLPAVDP 205
QY 138 -----SSRYLGSGLDFFLDITDTSVAKEKPARAPPE-----PHLFCAPG 179
DB 206 GKLSAASVSLASAGSAGAGLPGSKQVAPATPPPPPLAPVTPPPPLPABIFLBP 265
QY 180 SWE 182
DB 266 PME 268

RESULT 14

US-09-949-002-312
Sequence 312; Application US/09949002
Patent No. 6900016
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
TITLE OF INVENTION: WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
TITLE OF INVENTION: AND USES THEREOF
FILE REFERENCE: CLO00790
CURRENT APPLICATION NUMBER: US/09/949,002
CURRENT FILING DATE: 2000-01-28
PRIOR APPLICATION NUMBER: 60/231,401
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 10823
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 312
LENGTH: 824
TYPE: PRT
ORGANISM: Human
US-09-949-002-312

Query Match 8.8%; Score 90.5; DB 2; Length 824;
Best Local Similarity 25.0%; Pred. No. 3.9;
Matches 39; Conservative 24; Mismatches 42; Indels 51; Gaps 8;

QY 24 AGLQSYQOALRLISLDKVSRLGPRAPSLRRHVLINHTLOQALRL-----APAPALP 78
DB 673 AGIIVYRKRSRI-----LSRNAVAKTYMGRSNPLPFI-----QAASRVPAKGAAPAPSKG 722
QY 79 PEPLFLGEBDFLSATISILRELDTSMDGTSPONVTPPLGLONEVPPQDPVLEALS 138
DB 723 FOEL-----VPTTHPG-QPAKHPSASVALKR--PPAPPVTVSS-- 758
QY 139 SRVLGSDGLDFFLDITDTSVAKEK--PARAPPEPP 171
DB 759 -----PPFPVPTVTTQAPKQVIXKPTFAPEVPP 785

RESULT 15
US-08-887-518-2

Sequence 2, Application US/08887518
Patent No. 5843721
GENERAL INFORMATION:
APPLICANT: Rothe, Mike
APPLICANT: Wu, Lin
TITLE OF INVENTION: NIK Proteins, Nucleic Acids and Methods
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESS: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 268 BUSH STREET, SUITE 3200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/887,518
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: 797-008
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 343-4341
TELEFAX: (415) 343-4342
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 947 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-887-518-2

Query Match 8.8%; Score 90.5; DB 1; Length 947;
Best Local Similarity 25.8%; Pred. No. 4.7;
Matches 54; Conservative 17; Mismatches 87; Indels 51; Gaps 10;

QY 3 GGLKRSKSDLEBEERWMSPAQLQSYQOALRLISLDKVSRLGPRAPSLRRHVLINHTL 62
DB 659 GGLK---SPMRGEYKRPHPPPPNQANYQ-----TLHAQPRELSPRAPQPRPAESTTGRA 710
QY 63 QOQALALAPALPPE-----PLFLGEBDFLSATISILRELDTSMDGTSPONVTP 117
DB 711 PDLQ-----PLPPEPPPNKSPPLTLSEK-----SGMREPLPLSLBAPAPAPBS 758
QY 118 PLGLONEVPPQ-----DPPVLEALSRY-----LQDGLDFFLDITDTSVAKEKPA 164
DB 759 P-ERKATYPEQLEQLLELFLNLSQPFSLERQQLISCLSIDLSLSDS--EKNS 814
QY 165 RAPPEPPNLFCAPSW-----ENM 184
DB 815 KASGSBDTLSSGVHWSGQAEARSSWN 843

Search completed: February 6, 2006, 10:02:50
Job time : 50 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 6, 2006, 10:13:16 ; Search time 171 Seconds
(without alignments)
476.915 Million cell updates/sec

Title: US-10-069-386a-2

Perfect score: 1033

Sequence: 1 MEGSLKRRKHSDLREERERME.....AFGSWENELDHIMEIILGS 196

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA_Main:*

- 1: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep:*
- 3: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
- 4: /cgn2_6/ptodata/1/pubppaa/US10A_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/1/pubppaa/US10B_PUBCOMB.pep:*
- 6: /cgn2_6/ptodata/1/pubppaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	791	76.6	157	4	US-10-106-698-5516 Sequence 5516, Ap
2	645.5	62.5	142	4	US-10-029-386-33709 Sequence 33709, A
3	402	38.9	111	4	US-10-029-386-33708 Sequence 33708, A
4	217	21.0	236	6	US-11-095-870-15 Sequence 15, Appl
5	216	20.9	236	3	US-09-978-360A-568 Sequence 568, App
6	213	20.6	236	3	US-09-731-872-412 Sequence 412, App
7	213	20.6	236	3	US-09-876-997-412 Sequence 412, App
8	213	20.6	236	5	US-10-643-836-412 Sequence 412, App
9	213	20.6	236	6	US-11-095-870-14 Sequence 1065, Ap
10	213	20.6	236	4	US-10-029-386-34001 Sequence 34001, A
11	205	19.8	222	4	US-10-029-386-34001 Sequence 34001, A
12	136	13.2	237	4	US-10-076-069-2 Sequence 2, Appl1
13	125	12.1	314	4	US-10-170-385-31 Sequence 31, Appl
14	125	12.1	314	4	US-10-755-889-627 Sequence 627, Appl
15	125	12.1	314	6	US-11-095-870-12 Sequence 12, Appl
16	106	10.3	254	3	US-09-925-302-596 Sequence 596, App
17	106	10.3	254	3	US-09-925-302-596 Sequence 596, App
18	106	10.3	254	4	US-10-106-698-4456 Sequence 4456, App
19	105	10.2	241	4	US-10-076-069-4 Sequence 4, Appl1
20	105	10.2	241	5	US-10-965-898-51 Sequence 51, Appl
21	104.5	10.1	740	4	US-10-276-774-2083 Sequence 2083, Ap
22	104	10.1	322	5	US-10-450-763-53227 Sequence 53227, A
23	103.5	10.0	181	6	US-11-095-870-13 Sequence 13, Appl
24	98	9.5	562	4	US-10-104-047-3214 Sequence 3214, Ap
25	97	9.4	803	4	US-10-369-493-2536 Sequence 2536, Ap
26	97	9.4	803	5	US-10-732-923-7169 Sequence 7169, Ap
27	97	9.4	1099	4	US-10-425-115-195088 Sequence 195088, Ap

28	96	9.3	297	4	US-10-425-114-64002 Sequence 64002, A
29	96	9.3	817	4	US-10-369-493-1813 Sequence 1813, Ap
30	95.5	9.2	1228	4	US-10-437-963-188300 Sequence 188300, A
31	95	9.2	278	5	US-10-739-930-8565 Sequence 8565, Ap
32	95	9.2	713	4	US-10-437-963-137248 Sequence 137248, A
33	94	9.1	433	2	US-08-259-451-3 Sequence 3, Appl1
34	94	9.1	433	4	US-10-224-999A-3472 Sequence 3472, Ap
35	93.5	9.1	256	4	US-10-421-138A-312 Sequence 312, App
36	93.5	9.1	256	4	US-10-374-780A-1247 Sequence 1247, Ap
37	93.5	9.1	429	6	US-11-097-143-4629 Sequence 4629, Ap
38	93.5	9.1	663	4	US-10-755-889-4873 Sequence 4873, App
39	93.5	9.1	663	5	US-10-755-889-4873 Sequence 4873, App
40	93.5	9.1	1734	3	US-09-862-027-81 Sequence 81, Appl
41	93.5	9.1	1734	4	US-10-042-865-82 Sequence 82, Appl
42	93.5	9.1	1734	5	US-10-989-228-81 Sequence 81, Appl
43	93	9.0	157	4	US-10-767-701-54984 Sequence 54984, A
44	92.5	9.0	113	4	US-10-076-069-6 Sequence 6, Appl1
45	92	8.9	1103	4	US-10-437-963-166325 Sequence 166325, A

ALIGNMENTS

```

RESULT 1
US-10-106-698-5516
; Sequence 5516, Application US//10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: US/10/106,698
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: Patentin Ver. 3.0
; SEQ ID NO 5516
; LENGTH: 157
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (55)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC FEATURE
; LOCATION: (132)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-106-698-5516

Query Match
Best Local Similarity 76.6%; Score 791; DB 4; Length 157;
Matches 150; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 44 SGGPAPSLRRVLIHNTLQQLAALAPALPAPLPFLGEBDFSLSATGSLIRELD 103
A:|||||
DB 5 ALGPAPSLRRVLIHNTLQQLAALPAPLPFLGEBDFSLSATGSLIRELD 64
A:|||||
QY 104 TSMOTEPQNPVPTPLGIGNEVPPDPVFLAALSSRYLGSGLDFFLDITSAVEKRP 163
A:|||||
DB 65 TSMOTEPQNPVPTPLGIGNEVPPDPVFLAALSSRYLGSGLDFFLDITSAVEKRP 124
A:|||||
QY 164 ARAPEPHNLFCAFGSWENELDHIMEIILGS 196
A:|||||
DB 125 ARAPEPHNLFCAFGSWENELDHIMEIILGS 157
A:|||||

RESULT 2
US-10-029-386-33709

```


GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, Jean Baptiste
APPLICANT: Bouguerelet, Lydie
APPLICANT: Jozelet, Severin
TITLE OF INVENTION: FULL-LENGTH HUMAN CDNA ENCODING POTENTIALLY SECRETED PROTEINS
FILE REFERENCE: 78 USJ REG
CURRENT APPLICATION NUMBER: US/10/643,836
PRIOR FILING DATE: 2003-08-19
PRIOR APPLICATION NUMBER: US/09/731,872
PRIOR FILING DATE: 2000-12-07
PRIOR APPLICATION NUMBER: US 60/169,629
PRIOR FILING DATE: 1999-12-08
PRIOR APPLICATION NUMBER: US 60/187,470
PRIOR FILING DATE: 2000-03-06
NUMBER OF SEQ ID NOS: 482
SOFTWARE: Patent.pm
SEQ ID NO 412
LENGTH: 236
TYPE: PRT
ORGANISM: Homo sapiens
US-10-643-836-412

Query Match 20.6%; Score 213; DB 5; Length 236;
Best Local Similarity 32.6%; Pred. No. 5.1e-11;
Matches 76; Conservative 36; Mismatches 69; Indels 52; Gaps 13;

QY 4 GLKRGHSDLEEE---ERWESPAGLQSYQA-----LRLSLDKVRSGLPPAPS 51
DB 5 GLKRGHEEKEKEPLAVDSW-WLDGHAVAQAAPAVASSSLFDLSVLTCLHHSLQSSPD 63
QY 52 LRRHVLHNTLQOQLARLAPALPPEPL-----FLGEDPSLSTIGSILRE 101
DB 64 LRLVLVNTLRLRIQAS--MAPAALPVPSPAPASVADNMLASDALSMSMSLLED 121
QY 102 LDTSMDEPFPNPVTPLGLONEVPP-----QDPVFLRL--SSRYLGDGSLDDPFL 152
DB 122 L-SHIEGSLQAPQ-----LADGPPGRSIGGAAPSLGALDLAGPATGCLDDGSLGPE 175
QY 153 DDTSAVEKE---PARAPPEP-PHNLFCAPGSWE-----WNELDHIMEIILGS 196
DB 176 DDTSMYNEMLMAPASBGLKGPED---GPGKEAPBELDBALDVLMDVLVGT 225

RESULT 9
US-11-095-870-14
Sequence 14, Application US/11095870
Publication No. US2005022034A1
GENERAL INFORMATION:
APPLICANT: Hsu, Stephen I-Hong
TITLE OF INVENTION: MODULATION OF TRIP-BR FUNCTION AND METHOD OF TREATING
FILE REFERENCE: 93231-89
CURRENT APPLICATION NUMBER: US/11/095,870
PRIOR FILING DATE: 2005-03-31
PRIOR APPLICATION NUMBER: US 60/557,697
PRIOR FILING DATE: 2004-03-31
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin version 3.3
SEQ ID NO 14
LENGTH: 236
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: TRIP-Br1
US-11-095-870-14

Query Match 20.6%; Score 213; DB 6; Length 236;
Best Local Similarity 32.6%; Pred. No. 5.1e-11;
Matches 76; Conservative 36; Mismatches 69; Indels 52; Gaps 13;

QY 4 GLKRGHSDLEEE---ERWESPAGLQSYQA-----LRLSLDKVRSGLPPAPS 51
DB 5 GLKRGHEEKEKEPLAVDSW-WLDGHAVAQAAPAVASSSLFDLSVLTCLHHSLQSSPD 63
QY 52 LRRHVLHNTLQOQLARLAPALPPEPL-----FLGEDPSLSTIGSILRE 101
DB 64 LRLVLVNTLRLRIQAS--MAPAALPVPSPAPASVADNMLASDALSMSMSLLED 121
QY 102 LDTSMDEPFPNPVTPLGLONEVPP-----QDPVFLRL--SSRYLGDGSLDDPFL 152
DB 122 L-SHIEGSLQAPQ-----LADGPPGRSIGGAAPSLGALDLAGPATGCLDDGSLGPE 175
QY 153 DDTSAVEKE---PARAPPEP-PHNLFCAPGSWE-----WNELDHIMEIILGS 196
DB 176 DDTSMYNEMLMAPASBGLKGPED---GPGKEAPBELDBALDVLMDVLVGT 225

DB 5 GLKRGHEEKEKEPLAVDSW-WLDGHAVAQAAPAVASSSLFDLSVLTCLHHSLQSSPD 63
QY 52 LRRHVLHNTLQOQLARLAPALPPEPL-----FLGEDPSLSTIGSILRE 101
DB 64 LRLVLVNTLRLRIQAS--MAPAALPVPSPAPASVADNMLASDALSMSMSLLED 121
QY 102 LDTSMDEPFPNPVTPLGLONEVPP-----QDPVFLRL--SSRYLGDGSLDDPFL 152
DB 122 L-SHIEGSLQAPQ-----LADGPPGRSIGGAAPSLGALDLAGPATGCLDDGSLGPE 175
QY 153 DDTSAVEKE---PARAPPEP-PHNLFCAPGSWE-----WNELDHIMEIILGS 196
DB 176 DDTSMYNEMLMAPASBGLKGPED---GPGKEAPBELDBALDVLMDVLVGT 225

RESULT 10
US-10-296-115-1065
Sequence 1065, Application US/10296115
Publication No. US20040053248A1
GENERAL INFORMATION:
APPLICANT: Hysq Inc
TITLE OF INVENTION: No. US20040053248A1el Nucleic Acids and Polypeptides
FILE REFERENCE: 784PCT
CURRENT APPLICATION NUMBER: US/10/296,115
PRIOR FILING DATE: 2002-11-16
PRIOR APPLICATION NUMBER: US09/488,725
PRIOR FILING DATE: 2000-01-21
PRIOR APPLICATION NUMBER: US09/552,317
PRIOR FILING DATE: 2000-04-25
NUMBER OF SEQ ID NOS: 1478
SEQ ID NO 1065
LENGTH: 278
TYPE: PRT
ORGANISM: Homo sapiens
US-10-296-115-1065

Query Match 20.6%; Score 213; DB 4; Length 278;
Best Local Similarity 32.6%; Pred. No. 6.3e-11;
Matches 76; Conservative 36; Mismatches 69; Indels 52; Gaps 13;

QY 4 GLKRGHSDLEEE---ERWESPAGLQSYQA-----LRLSLDKVRSGLPPAPS 51
DB 47 GLKRGHEEKEKEPLAVDSW-WLDGHAVAQAAPAVASSSLFDLSVLTCLHHSLQSSPD 105
QY 52 LRRHVLHNTLQOQLARLAPALPPEPL-----FLGEDPSLSTIGSILRE 101
DB 106 LRLVLVNTLRLRIQAS--MAPAALPVPSPAPASVADNMLASDALSMSMSLLED 163
QY 102 LDTSMDEPFPNPVTPLGLONEVPP-----QDPVFLRL--SSRYLGDGSLDDPFL 152
DB 164 L-SHIEGSLQAPQ-----LADGPPGRSIGGAAPSLGALDLAGPATGCLDDGSLGPE 217
QY 153 DDTSAVEKE---PARAPPEP-PHNLFCAPGSWE-----WNELDHIMEIILGS 196
DB 218 DDTSMYNEMLMAPASBGLKGPED---GPGKEAPBELDBALDVLMDVLVGT 267

RESULT 11
US-10-029-386-34001
Sequence 34001, Application US/10029386
Publication No. US20030194704A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
TITLE OF INVENTION: HUMAN GENE--DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR C
FILE REFERENCE: AEOICA-X-2
CURRENT APPLICATION NUMBER: US/10/029,386
PRIOR FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 34288
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 34001

LENGTH: 222
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC010271.5
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.46
OTHER INFORMATION: SWISSPROT HIT: Q14140, EVALU2.70e-01
US-10-029-386-34001

Query Match 19.8%; Score 205; DB 4; Length 222;
Best Local Similarity 32.6%; Pred. No. 2.5e-10;
Matches 75; Conservative 34; Mismatches 69; Indels 52; Gaps 13;

QY 4 GIKRKHSDLEEEER-ERMWSPAGLSYQQA-----LIRSLDKVQSLGRAPRS 51
DB 5 GIKRKRREERKEPLAVNSW-WLDGHTAVNAAPRAVSSSLFDLSVTKLHHSLQSSPD 63
QY 52 LRRHVLHNTLQOLQALRLAPALPPEPL-----FLGEEDPSLSATIGSLIRE 101
DB 64 LRHLVAVNTLRLRIQAS--MAPAALPVPSPPAAPSVADNLLASSDALSASMASLIED 121
QY 102 LDTSMDCHEPPONPTPLGLONEVPP-----QDDPVFLAL--SKRYLGSGLDDPFL 152
DB 122 L-SHIEGSLQAPQP-----LADEGPPGRSIGGAAPSLGALDLGPATGCLDDGLEGLFE 175
QY 153 DIDTSAVEKE--PARAPPEP-PHNLFCAPGSWE-----WNELDHIMEII 193
DB 176 DIDTSMDNELMAPASEGKPGPED--GPGKEAPELDEALDYLMDVL 222

RESULT 12

US-10-076-069-2
Sequence 2, Application US/10076069
Publication No. US20020177214A1
GENERAL INFORMATION:
APPLICANT: NACHTMAN, RONALD
TITLE OF INVENTION: HEPB, A NOVEL GENE WITH A ROLE IN HEMATOPOIETIC AND NEURAL DEVELOPMENT
FILE REFERENCE: 39532-176599
CURRENT FILING DATE: 2002-02-15
PRIOR APPLICATION NUMBER: US 60/268,923
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2
LENGTH: 237
TYPE: PRT
ORGANISM: Mus musculus
US-10-076-069-2

Query Match 13.2%; Score 136; DB 4; Length 237;
Best Local Similarity 26.1%; Pred. No. 0.00049;
Matches 63; Conservative 32; Mismatches 86; Indels 60; Gaps 9;

QY 4 GIKRKHSDLEEEERWEMSPA-GLOSXQALLRISLDKQKSLGRAPSLRHVLIHNTL 62
DB 5 GIKRKYGOEBEGEVGTVPYSYLSQ--RQSLDMSLVKLQCHMLVEPPLCRSVLIANTV 62
QY 63 QOLQALR-----LAP--APALPPEPLFLGEEDPSLSATIGSLIRELDTSMDCHEPPQ 113
DB 63 ROIQESMGODGWGMAAPONVDAVERLVSTE--ILCRTRGABEEHPAPAELEDAPIQ 119
QY 114 NPVTPLGLONEVPPQDP-----VFLALSSRYLGDSDGLDDP 151
DB 120 NNSVSELPYGSAPGGRNPQSSLSLWMDSPQENKSGFQSLDQJFETLENK--NSSVBELEF 177
QY 152 LIDITSAAVEKEPAR-----APPEPPHNLFCAPGSWENELDHIMEII 193
DB 178 SDVDSYDLDITVLTMGMSGTKSLCNGLEGFAAATPPSSSTCKS---DLALDHVLEII 234
QY 194 L 194

DB 235 V 235

RESULT 13

US-10-170-385-31
Sequence 31, Application US/10170385
Publication No. US20030203372A1
GENERAL INFORMATION:
APPLICANT: Ward, Neil Raymond
APPLICANT: Mundy, Christopher Robert
APPLICANT: Kan, On
APPLICANT: Harris, Robert Alan
APPLICANT: White, Jonathan
APPLICANT: Binley, Katie Mary
APPLICANT: Rayner, William Nigel
APPLICANT: Naylor, Stuart
APPLICANT: Kingman, Susan Mary
APPLICANT: Krige, David
TITLE OF INVENTION: ANALYSIS METHOD
FILE REFERENCE: 532682000100
CURRENT APPLICATION NUMBER: US/10/170,385
PRIOR FILING DATE: 2002-06-12
PRIOR APPLICATION NUMBER: PCT/GB02/01662
PRIOR FILING DATE: 2002-04-08
PRIOR APPLICATION NUMBER: PCT/GB01/05458
PRIOR FILING DATE: 2001-12-10
NUMBER OF SEQ ID NOS: 549
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 31
LENGTH: 314
TYPE: PRT
ORGANISM: Homo Sapiens
US-10-170-385-31

Query Match 12.1%; Score 125; DB 4; Length 314;
Best Local Similarity 22.7%; Pred. No. 0.007;
Matches 71; Conservative 34; Mismatches 88; Indels 120; Gaps 12;

QY 2 EGKIKRKHSDLEEEERWEMSPAQS-----YQALLRISLDKQKSLGRAPSLRH 55
DB 4 KGG-KRKHDEHDDGLEKIVSPCDGPSKVSYTLQKQTFINISLMKLVNHRPLTEPSLOKT 62
QY 56 VLIHNTLQOLQALR-----PALPE--PLFLGEEDPSLSATIGSLIRELD 103
DB 63 VLIHNTLQOLQALR-----PALPE--PLFLGEEDPSLSATIGSLIRELD 103
QY 71 -----LAP-----PALPE--PLFLGEEDPSLSATIGSLIRELD 103
DB 123 TTPLEACLTPLASLEDDDDTCTSQAMQPTAATKLSPPALLPEKD--SFSSALDEIEBLCP 181
QY 104 TSM-----GTEPPONPTPLGLONEVPPQ-----PDDVFLAL-----SSRYLGD 145
DB 182 TSTSTEAATTAATSVKGTSSSEAGTYKLDGPESRADDKMLDNGNFEITTSSTPLTDL 241
QY 146 GLDD-FPLDITSAVEKEP-----ARAPPEPPHNLF-----CAPGSWE 183
DB 242 TIDDLTLPADITSMYDPPCTSSSGTASAKMAPVADDLTKTLAPYSQPTPSQEPKNDL 301
QY 184 NELDHIMEIIGS 196
DB 302 TELDHIMEIVIGS 314

RESULT 14

US-10-755-889-627
Sequence 627, Application US/10755889
Publication No. US20040171823A1
GENERAL INFORMATION:
APPLICANT: Bristol-Myers Squibb Company
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-kB
FILE REFERENCE: D0284 NP
CURRENT APPLICATION NUMBER: US/10/755,889

```

CURRENT FILING DATE: 2004-01-13
PRIOR APPLICATION NUMBER: U.S. 60/440,068
PRIOR FILING DATE: 2003-01-14
PRIOR APPLICATION NUMBER: U.S. 60/465,757
PRIOR FILING DATE: 2003-05-12
NUMBER OF SEQ ID NOS: 823
SOFTWARE: PatentIn version 3.2
SEQ ID NO 627
LENGTH: 314
TYPE: prt
ORGANISM: Homo sapiens
US-10-755-889-627

```

Query Match	12.1%;	Score 125;	DB 4;	Length 314;
Best Local Similarity	22.7%;	Pred. No. 0.007;		
Matches	71;	Mismatches	88;	Indels 120;
		Gaps	12;	

```

Oy      2 EGGLKRXKSDLEBBEREMWSPAGLOS-----YQALRLISUDYQXRLSPAPSLRRI 55
      4 KKG-RRKDEHNDGEGKIVSPCDGPKSVSYTLQRTIFINISLMKLYNHRPLTEPSLOKT 62
Oy      56 VLINHTLOOLQALR-----70
      63 VLINMLKRIQELKQEGSLRPMFTPSQPTTBSPDSTYERAPAFSHLASPSHPCDLGS 122
Oy      71 -----LAPA-----PALPPE-----PLFGEHDELSLATTIGSLIRELD 103
Db      123 TTPLEACILTPALLEDDDDTFCTSQAMQTPATPKLSPPALLPCKD-SFSALDEIBELCP 181
Oy      104 TSMQ-----GTEPQNPVTPYGLONEVPPQ--PDVFFLEAL-----SSRYIGDS 145
Db      182 TSTSEATPAARDVSYGTSSSAGTQQLDGPQESRADDSKLMBSLPCGNFEITTTGTGFTLD 241
Oy      146 GLDD-FLPDDIDTSSAVEKER-----AAPPEPPHNLF-----CAPGSMEW 183
Db      242 TLDDTLFLADIDTSMWDFDPCSTSSSGTASQAPVASADLLKTLAPYSSQPYTPSQPFKMDL 301
Oy      184 NELDIMEITIGS 196
Db      302 TELDHIMEIVLGS 314

```

```

RESULT 15
US-11-095-870-12
: Sequence 12, Application US/11095870
: Publication No. US20050222034A1
: GENERAL INFORMATION:
: APPLICANT: Hsu, Stephen I-Hong
: TITLE OF INVENTION: MODULATION OF TRIP-BR FUNCTION AND METHOD OF TREATING
: TITLE OF INVENTION: MODULATION OF TRIP-BR FUNCTION AND METHOD OF TREATING
: FILE REFERENCE: 93231-89
: CURRENT APPLICATION NUMBER: US/11/095,870
: CURRENT FILING DATE: 2005-03-31
: PRIOR APPLICATION NUMBER: US 60/557,697
: PRIOR FILING DATE: 2004-03-31
: NUMBER OF SEQ ID NOS: 16
: SOFTWARE: PatentIn version 3.3
: SEQ ID NO 12
: LENGTH: 314
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: OTHER INFORMATION: TRIP-BR2
US-11-095-870-12

```

Query Match	12.1%;	Score 125;	DB 6;	Length 314;
Best Local Similarity	22.7%;	Pred. No. 0.007;		
Matches	71;	Conservative	34;	Mismatches 88;
				Indels 120;
				Gaps 12

```
Oy      2 EGGLEKSHDLEBBEREWSPALQS-----YQALLRLISDKVQSLGPRAPSLRRH    55
          |||||:::||::|||::|||::|||::|||::|||::|||::|||::|||::
Db      4 KGG-KKKFDEHEGLBGKI VSPDGPSKVSYTLQQTIFNISLMLKNHRPLTSPSLQKT    62
```

```

OY  VLIHHTLQOLAAAR-----70
    |||::|:|
Db  VLIINMARIQEBELKQBSGLRPMFTPSQPTTERPSDYREAPAFSHLSPSSHPCDLGS 122
63  VLIINMARIQEBELKQBSGLRPMFTPSQPTTERPSDYREAPAFSHLSPSSHPCDLGS 122
OY  71-----LAP-----PALPPE--PLFLGHEPDSLANTISSILARELD 103
Db  123  TTPLEACTLPASLLEDDDDTCTSQAMOPTAFTKLSPPALLPEKD--SFSALDISELCP 181
OY  104  TSMD-----GTEPQNPPTPLGLONEVPPQ-----PDVFLLEAL-----SSRYLGDS 145
Db  182  TSTSEATATATDSYKCTSSSEACTQKLDOPQBSRADSLKMSLSPNFEITTSGTGFTDL 241
OY  146  GIDD--FLDIDTSSAVEKEP-----ARAPPEPPHNLF-----CAPGSNEW 183
Db  242  TLDDTLFLFDIDTSMYDFPCTSSSGTASKMAVSDLDLKTLLAPYSQPYTPSQPFKMDL 301
OY  184  NELDIMEILIGS 196
Db  302  TELDHIMEVLVGS 314

```

Search completed: February 6, 2006, 10:16:34
Job time : 172 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 6, 2006, 10:13:51 ; Search time 17 Seconds
(without alignments)
135.102 Million cell updates/sec

Title: US-10-069-386a-2

Perfect score: 1033
Sequence: 1 MEGUKRKHSDEEEERWE.....ADGSWENLDHIMEIILGS 196

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 88029 seqs, 11718060 residues

Total number of hits satisfying chosen parameters: 88029

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA_New:*
1: /cgn2_6/protdata/1/pubppaa/US08_NEW_PUB.pep:*
2: /cgn2_6/protdata/1/pubppaa/US06_NEW_PUB.pep:*
3: /cgn2_6/protdata/1/pubppaa/US07_NEW_PUB.pep:*
4: /cgn2_6/protdata/1/pubppaa/PCT_NEW_PUB.pep:*
5: /cgn2_6/protdata/1/pubppaa/US09_NEW_PUB.pep:*
6: /cgn2_6/protdata/1/pubppaa/US10_NEW_PUB.pep:*
7: /cgn2_6/protdata/1/pubppaa/US11_NEW_PUB.pep:*
8: /cgn2_6/protdata/1/pubppaa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	216	20.9	236	5	US-09-978-360A-568 Sequence 568, App
2	94	9.1	372	6	US-10-714-887-114 Sequence 114, App
3	90.5	8.8	324	6	US-10-511-314-19 Sequence 19, App1
4	90.5	8.8	324	6	US-10-511-722-19 Sequence 19, App1
5	88.5	8.6	1122	6	US-10-821-234-1657 Sequence 1657, App
6	88.5	8.6	1122	6	US-10-964-313-4 Sequence 4, App1
7	84	8.1	1015	7	US-11-169-041-217 Sequence 217, App
8	83.5	8.1	1823	6	US-10-995-561-988 Sequence 988, App
9	83.5	8.1	2102	6	US-10-995-561-990 Sequence 990, App
10	83.5	8.1	2108	6	US-10-995-561-989 Sequence 989, App
11	83.5	8.1	2157	6	US-10-995-561-991 Sequence 991, App
12	82	7.9	2101	6	US-10-857-780-23 Sequence 23, App1
13	81.5	7.9	384	7	US-11-134-563-2 Sequence 2, App1
14	81	7.8	891	7	US-11-205-109-18 Sequence 18, App1
15	80.5	7.8	606	6	US-10-055-877-58 Sequence 58, App1
16	80.5	7.8	1532	6	US-10-821-234-914 Sequence 914, App1
17	80.5	7.8	7968	7	US-11-143-980-49 Sequence 49, App1
18	79	7.6	625	7	US-11-143-984A-39 Sequence 39, App1
19	79	7.6	625	7	US-11-143-984A-110 Sequence 110, App
20	79	7.6	718	6	US-10-918-857-2 Sequence 2, App1
21	79	7.6	790	6	US-10-818-857-6 Sequence 6, App1
22	78.5	7.6	261	7	US-11-169-041-190 Sequence 190, App
23	78.5	7.6	1377	6	US-10-821-234-1070 Sequence 1070, App
24	77	7.5	558	6	US-10-504-364-3 Sequence 3, App1
25	77	7.5	558	6	US-10-504-364-4 Sequence 4, App1

26	77	7.5	1560	7	US-11-059-982-1 Sequence 1, App1
27	76.5	7.4	616	6	US-10-982-545-5 Sequence 5, App1
28	76.5	7.4	753	7	US-11-037-243-68 Sequence 68, App1
29	76.5	7.4	1041	6	US-10-995-561-780 Sequence 780, App
30	76.5	7.4	1041	6	US-10-995-561-782 Sequence 782, App
31	76.5	7.4	1097	6	US-10-995-561-781 Sequence 781, App
32	76	7.4	863	7	US-11-169-041-167 Sequence 167, App
33	75.5	7.3	304	7	US-11-134-563-4 Sequence 4, App1
34	75	7.3	559	7	US-11-150-945-38 Sequence 38, App1
35	75	7.3	559	7	US-11-149-945-3 Sequence 3, App1
36	75	7.3	672	7	US-11-000-463-455 Sequence 455, App
37	74.5	7.2	571	7	US-11-121-438-12 Sequence 12, App1
38	74.5	7.2	884	6	US-10-995-561-786 Sequence 786, App
39	74.5	7.2	2630	7	US-11-186-731-2 Sequence 2, App1
40	74.5	7.2	7968	7	US-11-186-731-5 Sequence 5, App1
41	74	7.2	578	7	US-11-037-243-100 Sequence 100, App
42	74	7.2	746	6	US-10-828-831-5 Sequence 5, App1
43	74	7.2	746	6	US-10-828-831-7 Sequence 7, App1
44	74	7.2	1041	6	US-10-828-831-9 Sequence 9, App1
45	74	7.2	1085	6	US-10-523-477-13 Sequence 13, App1

ALIGNMENTS

RESULT 1
US-09-978-360A-568
; Sequence 568, Application US/09978360A
; Publication No. US20060096633A9
; GENERAL INFORMATION:
; APPLICANT: Edwards, Jean-Baptiste Dumas Milne
; APPLICANT: Duclert, Aymeric
; APPLICANT: Bouquelier, Lydie
; APPLICANT: Joberet, Severin
; APPLICANT: Clineel, Catherine
; TITLE OF INVENTION: Complementary DNA's Encoding Proteins with Signal Peptides
; FILE REFERENCE: 56, US4, CIP
; CURRENT APPLICATION NUMBER: US/09/978,360A
; PRIOR FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: US 60/066,677
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: US 60/069,957
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: US 60/074,121
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: US 60/081,563
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: US 60/096,116
; PRIOR FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: US 60/099,273
; PRIOR FILING DATE: -09-04
; PRIOR APPLICATION NUMBER: US 09/191,997
; PRIOR FILING DATE: 1998-11-13
; PRIOR APPLICATION NUMBER: US 09/215,435
; PRIOR FILING DATE: 1998-12-17
; PRIOR APPLICATION NUMBER: PCT/IB98/02122
; PRIOR FILING DATE: 1998-12-17
; PRIOR APPLICATION NUMBER: US 09/247,155
; PRIOR FILING DATE: 1999-02-09
; Remaining Prior Application data removed - See file wrapper or PAM.
; NUMBER OF SEQ ID NOS: 810
; SOFTWARE: Patent.pm
; SEQ ID NO 568
; LENGTH: 236
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -31..-1
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (28, 30, 40, 67, 86, 117, 120)
; OTHER INFORMATION: unknown

US-09-978-360A-568

Query Match 20.9%; Score 216; DB 5; Length 236;
 Best Local Similarity 32.6%; Pred. No. 4,3e-13;
 Matches 76; Conservative 36; Mismatches 69; Indels 52; Gaps 13;

QY 4 GIKRHSOLEEER---EREMSPAGLOSIOQA-----LIRISLDKVRSLGPRAPS 51
 DB 5 GIKRKEEERKEEPLAVDSW-WLDPGHAAVQAAPVAVSSSLFDSISVTKLHSIQXSPD 63
 QY 52 LRRHVLHINTLQOQLALRLAPALPPEPL-----PLGDEDSLSATISILRE 101
 DB 64 LRLHLVAVKTLKRIOQS--MAPAALPPVPTPPAPXVADNLLASSDALSSMAXLLED 121
 QY 102 LDTSMGTEPPQNPVPLGLONEVPPDPD-----PVLEAL-----SSRYLGDSGLDPEFL 152
 DB 122 L-SHIEGEGSQAPQP-----LADGPPGRIIGXPPYIGALDLGPAVTCGLDNGEGJPR 175
 QY 153 DITSAVEKE---PARAPPEP-PHNLFCAPGSWE-----WNELDHIMEILLGS 196
 DB 176 DITSMYDNELWAPASBGKPGPED---GPGKEAPELDEARELDYIMDLVGT 225

RESULT 2

US-10-714-887-114
 ; Sequence 114, Application US/10714887
 ; Publication No. US20060015972A1
 ; GENERAL INFORMATION:

APPLICANT: Mendel Biotechnology, Inc.
 APPLICANT: HEARD, Jacqueline
 APPLICANT: RICHMANN, Jose Luis
 APPLICANT: CREBIMAN, Robert
 APPLICANT: RATCLIFFE, Oliver
 APPLICANT: CANALES, Roger
 APPLICANT: REPETTI, Peter
 APPLICANT: KUMIMOTO, Roderick W
 APPLICANT: GUTTERSON, Neal
 APPLICANT: REUBER, T. Lynne
 APPLICANT: PINEDA, Omeira
 APPLICANT: SHERMAN, Bradley K
 TITLE OF INVENTION: PLANT TRANSCRIPTIONAL REGULATORS OF DROUGHT STRESS
 FILE REFERENCE: MB10058-CIP
 CURRENT APPLICATION NUMBER: US/10/714,887
 PRIOR FILING DATE: 2003-11-13
 PRIOR APPLICATION NUMBER: 10/412,699
 PRIOR FILING DATE: 2003-04-10
 PRIOR APPLICATION NUMBER: 09/506,720
 PRIOR FILING DATE: 2000-02-17
 PRIOR APPLICATION NUMBER: 60/135,134
 PRIOR FILING DATE: 1999-05-20
 PRIOR APPLICATION NUMBER: 09/394,519
 PRIOR FILING DATE: 1999-09-13
 PRIOR APPLICATION NUMBER: 09/533,392
 PRIOR FILING DATE: 2000-03-22
 PRIOR APPLICATION NUMBER: 09/533,029
 PRIOR FILING DATE: 2000-03-22
 PRIOR APPLICATION NUMBER: 09/532,591
 PRIOR FILING DATE: 2000-03-22
 PRIOR APPLICATION NUMBER: 09/533,030
 PRIOR FILING DATE: 2000-03-22
 PRIOR APPLICATION NUMBER: 60/125,814
 PRIOR FILING DATE: 1999-03-23
 PRIOR APPLICATION NUMBER: 09/713,994
 PRIOR FILING DATE: 2000-11-16
 Remaining prior Application data removed - See File Wrapper or PALM.
 NUMBER OF SEQ ID NOS: 430
 SOFTWARE: PatentIn version 3.2
 SEQ ID NO 114
 LENGTH: 372
 TYPE: PRT
 ORGANISM: Zea mays
 FEATURE:
 OTHER INFORMATION: G3680 polypeptide Orthologous to G2999

US-10-714-887-114

Query Match 9.1%; Score 94; DB 6; Length 372;
 Best Local Similarity 33.8%; Pred. No. 0.14;
 Matches 27; Conservative 11; Mismatches 32; Indels 10; Gaps 3;

QY 105 SMDGTEPPQNPVPLGLONEVPP-----OPDPVLEKLSRSRYLGDSGLDPEFLDITSA 158
 DB 87 TLGSGPPPPAP-APLALP--PPSVWGHQPHRAREETPRHHFQVADADSDSDSGSE 142
 QY 159 VEKEPARAPPEPPHNLFCAP 178
 DB 143 YDERRSVSPPPPHHLPLAP 162

RESULT 3

US-10-511-314-19
 ; Sequence 19, Application US/10511314
 ; Publication No. US20050272633A1
 ; GENERAL INFORMATION:

APPLICANT: Yeda Research and Development Co. Ltd.
 APPLICANT: Wallach, David
 APPLICANT: Shmushkovich, Taisia
 APPLICANT: Ramakrishnan, Parameswaran
 TITLE OF INVENTION: Derivatives of NIK, their production and use
 FILE REFERENCE: 814
 CURRENT APPLICATION NUMBER: US/10/511,314
 PRIOR FILING DATE: 2004-10-15
 PRIOR APPLICATION NUMBER: 149217
 PRIOR FILING DATE: 2002-04-18
 PRIOR APPLICATION NUMBER: 152183
 PRIOR FILING DATE: 2002-10-08
 NUMBER OF SEQ ID NOS: 19
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 19
 LENGTH: 324
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-511-314-19

Query Match 8.8%; Score 90.5; DB 6; Length 324;
 Best Local Similarity 25.8%; Pred. No. 0.24;
 Matches 54; Conservative 17; Mismatches 87; Indels 51; Gaps 10;

QY 3 GGIKRRHSDLSEEREMWSPAGLOSIOQALRLISLDKVRSLGPRAPSLRHVLIHNTL 62
 DB 36 GGLK---SPWRGEYKEPRHPNPQANYHQ-----TLHAQPRRLSPRAPGPRRAETTGRA 87
 QY 63 QOQLAALRLAPALPPE-----PLFGEHEDSLSATISILRELDTSMDGTEPPQNPVT 117
 DB 88 PDLQ-----PPLPPEPPBNKSPYTLISKEE-----SGMWEPLPLSLLEAPAPAPNPS 135
 QY 118 PLGLONEVPPQ-----PDPVLEALSRY-----LDGSLDPEFLDITSAVEKEPA 164
 DB 136 P-ERKATVPDEQLQOLELIEFLNLSQSPSLERQELISLSDLSLSDS---EKNS 191
 QY 165 RAPPEPPHNLFCAPGSW-----EWN 184
 DB 192 KASQSSRDTLSSGCHWSQGAARASSWN 220

RESULT 4

US-10-511-722-19
 ; Sequence 19, Application US/10511722
 ; Publication No. US20050287144A1
 ; GENERAL INFORMATION:

APPLICANT: Yeda Research and Development Co. Ltd.
 APPLICANT: Wallach, David
 APPLICANT: Shmushkovich, Taisia
 APPLICANT: Ramakrishnan, Parameswaran
 TITLE OF INVENTION: Derivatives of the IL-2 receptor Gamma chain, their preparation
 FILE REFERENCE: 530
 CURRENT APPLICATION NUMBER: US/10/511,722

;; CURRENT FILING DATE: 2004-10-18
;; PRIOR APPLICATION NUMBER: 149217
;; PRIOR FILING DATE: 2002-04-18
;; PRIOR APPLICATION NUMBER: 152183
;; PRIOR FILING DATE: 2002-10-08
;; NUMBER OF SEQ ID NOS: 21
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 19
;; LENGTH: 324
;; TYPE: PRF
;; ORGANISM: Homo sapiens
US-10-511-722-19

Query Match 8.6%; Score 90.5; DB 6; Length 324;
Best Local Similarity 25.8%; Pred. No. 0.24;
Matches 54; Conservative 17; Mismatches 87; Indels 51; Gaps 10;

QY 3 GGLKRRHSDLEEEERWMSPAQLSYOQALLRISLDKVRSLGPPAPSLRHVLIHNTL 62
DB 36 GGLK---SPWRGEYKPRHPPNQNAYHQ-----TLHAQPRRLSPRAPPPRAEETTGRA 87
QY 63 QQLQALRLAPALPPE-----PLFLGEBDFSLKATIGSILRELDTSMDGTEPPQNPPT 117
DB 88 PDLQ-----PPLPPEPPEPNKSPPLTLSEK-----SGWMEPLPLSLBPAPAPNPSS 135
QY 118 PLGLNEVPPO-----PDVPLEALSSRY-----LGDGGLDDEFLDIDTSAVEKEPA 164
DB 136 P-ERKATYPEQLQOLETELFLNSLSQPSLEQEQILSCISLISLSDS---EKNDP 191
QY 165 RAPPEPHNLFCAPGSW-----EMN 184
DB 192 KASQSSRDTLSSGVHSMSSQAEARSSSN 220

RESULT 5
US-10-821-234-1657
;; Sequence 1657, Application US/10821234
;; Publication No. US20050255114A1
;; GENERAL INFORMATION:
;; APPLICANT: Labat, Ivan
;; APPLICANT: Stache-Crain, Birgit
;; APPLICANT: Andarmani, Sueann
;; APPLICANT: Tang, Y. Tom
;; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
;; FILE REFERENCE: 821A
;; CURRENT APPLICATION NUMBER: US/10/821,234
;; CURRENT FILING DATE: 2004-04-07
;; PRIOR APPLICATION NUMBER: US 60/462,047
;; PRIOR FILING DATE: 2003-04-07
;; NUMBER OF SEQ ID NOS: 1704
;; SOFTWARE: PC_SEQ_genes Version 1.0
;; SEQ ID NO 1657
;; LENGTH: 1122
;; TYPE: PRF
;; ORGANISM: Homo sapiens
US-10-821-234-1657

Query Match 8.6%; Score 88.5; DB 6; Length 1122;
Best Local Similarity 24.4%; Pred. No. 1.7;
Matches 50; Conservative 23; Mismatches 75; Indels 57; Gaps 9;

QY 13 EEEERWMSPAQLSYOQ-----ALLRISLDKVRSLGPPAP-----SL 52
DB 396 QOEAEAR-----QALQSLRQGGTLTGKFMSTSSIPGCLLGVALE-----GDSPHGHASL 444
QY 53 RHHVLIHNTLQQLQALRLAPALPPEPLFLGEBDFSLKATIGSILRELDTSMDGTEPP 112
DB 445 LQHVLL---LEQARQOSTLIAVPLHGQSPVTGERVATSMRTVGKLPRIKPLSRQSSP- 500
QY 113 QNPVTPPLGLONEVPPQPDVPLEALSSR--YLGDGGLDDEFLDIDTSAVEKEPARAPPEP 170
DB 501 -LPQSPQALQQLVWMOQOHOQFLKQKQQLQGLK-----ILTKTGELPRQPTTHPEBT 552

QY 171 PHNLFCAQGSWEMNLDHIMEIILG 195
DB 553 E-----EBLTREQQEVILG 565

RESULT 6
US-10-964-313-4
;; Sequence 4, Application US/10964313
;; Publication No. US20050287629A1
;; GENERAL INFORMATION:
;; APPLICANT: GROZINGER, CHRISTINA M.
;; APPLICANT: HASSIG, CHRISTIAN A.
;; APPLICANT: SCHREIBER, STUART L.
;; TITLE OF INVENTION: CLASS II HUMAN HISTONE DEACETYLASES, AND USES RELATED
;; TITLE OF INVENTION: THERETO
;; FILE REFERENCE: HIV-037.02
;; CURRENT APPLICATION NUMBER: US/10/964,313
;; CURRENT FILING DATE: 2004-10-13
;; PRIOR APPLICATION NUMBER: 09/800,187
;; PRIOR FILING DATE: 2001-03-05
;; PRIOR APPLICATION NUMBER: 60/186,802
;; PRIOR FILING DATE: 2000-03-03
;; NUMBER OF SEQ ID NOS: 73
;; SOFTWARE: PatentIn Ver. 3.3
;; SEQ ID NO 4
;; LENGTH: 1122
;; TYPE: PRF
;; ORGANISM: Homo sapiens
US-10-964-313-4

Query Match 8.6%; Score 88.5; DB 6; Length 1122;
Best Local Similarity 24.4%; Pred. No. 1.7;
Matches 50; Conservative 23; Mismatches 75; Indels 57; Gaps 9;

QY 13 EEEERWMSPAQLSYOQ-----ALLRISLDKVRSLGPPAP-----SL 52
DB 396 QOEAEAR-----QALQSLRQGGTLTGKFMSTSSIPGCLLGVALE-----GDSPHGHASL 444
QY 53 RHHVLIHNTLQQLQALRLAPALPPEPLFLGEBDFSLKATIGSILRELDTSMDGTEPP 112
DB 445 LQHVLL---LEQARQOSTLIAVPLHGQSPVTGERVATSMRTVGKLPRIKPLSRQSSP- 500
QY 113 QNPVTPPLGLONEVPPQPDVPLEALSSR--YLGDGGLDDEFLDIDTSAVEKEPARAPPEP 170
DB 501 -LPQSPQALQQLVWMOQOHOQFLKQKQQLQGLK-----ILTKTGELPRQPTTHPEBT 552
QY 171 PHNLFCAQGSWEMNLDHIMEIILG 195
DB 553 E-----EBLTREQQEVILG 565

RESULT 7
US-11-169-041-217
;; Sequence 217, Application US/11169041
;; Publication No. US20060019284A1
;; GENERAL INFORMATION:
;; APPLICANT: Bristol-Myers Squibb Company
;; TITLE OF INVENTION: IDENTIFICATION OF POLYNUCLEOTIDES FOR PREDICTING ACTIVITY OF
;; TITLE OF INVENTION: COMPOUNDS THAT INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE
;; TITLE OF INVENTION: KINASES AND/OR PROTEIN TYROSINE KINASE PATHWAYS IN LUNG CANCER
;; TITLE OF INVENTION: CELLS
;; FILE REFERENCE: 10001 NP
;; CURRENT APPLICATION NUMBER: US/11/169,041
;; CURRENT FILING DATE: 2005-06-28
;; PRIOR APPLICATION NUMBER: 60/584,405
;; PRIOR FILING DATE: 2004-06-30
;; NUMBER OF SEQ ID NOS: 527
;; SOFTWARE: PatentIn version 3.2
;; SEQ ID NO 217
;; LENGTH: 1015
;; TYPE: PRF
;; ORGANISM: Homo sapiens
US-11-169-041-217

Query Match	8.1%;	Score 84;	DB 7;	Length 1015;
Best Local Similarity	25.7%;	Pred. No. 3.9;		
Matches	53;	Conservative	20;	Mismatches 63;
			Indels	70;
			Gaps	10;

```

OY      18 WWMENPAGIGSYQOALLTRIS-----LDKVRSL-----GPRAPS- 51
Db      71 RYEVSPVALQRLRVALLQKLSGCGFTWQDDYDTQYWMQGLADLPKTYLRREPAASPARPSK 130
OY      52 -----LRRHVLIHNTLQOLQALRIAPALPPEELFTGEEDFSL 91
Db      131 HSYGSERRYSREGAALANALRRRLPFLBALSQAPASDVLAFTHTAQDPRPAAGDRFSS- 169
OY      92 SATISIIIRLBDTSMGDTFRPQNPVTPPLQONEVP-----QPDVFLEALSRLIGDS 145
Db      190 -----ESIIITYVAHTSALTYPP-GPRQO--LRBDLLPRTLGQLQPD-----ELSPKV--DS 235
OY      146 GLDDFFLDDIDTSAVEKEPARAPPEPP 171
Db      236 GVDRRHLL---MAALSTAYAAQRRPPAP 258

```

RESULT 8
US-10-995-561-988
; Sequence 988, Application US/10995561
; Publication No. US20050272054A1

```

; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: C1001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 988
; LENGTH: 1823
; TYPE: PRM
; ORGANISM: Homo sapiens
US-10-995-561-988

```

Query Match	8.1%;	Score 83.5;	DB 6;	Length 1823;
Best Local Similarity	20.7%;	Pred. No. 8.7;		
Matches	45;	Conservative	18;	Mismatches 63;
				Indels 91;
				Gaps 9

```
QY      95 IGST-----LRELDTSMDGTBPONPV-T-PLGLONEVPRQDPVFLG----- 135
Db      353 LGAVPAPQAPPPPKKALYPGALGRPPPMFPMMFDPRMM--TTPYVDPRLLQGRPELDFX 410
QY      136 -----ALSRYLGDSG-----LDDEF----- 151
Db      411 PPGVHPSPGVLPFRSRSSGGSSSEPFDRHAPAMLREGCTPPVPDKLAMVGDFATPAEPR 470
QY      152 -----LDIDTAIVEKEPARAPDEPPNHLFCARG 179
Db      471 PUTSPLRQAADSDCKMGRSETTPVPPPPY-LASLYRG 506
```

```

RESULT 9
US-10-995-561-990
; Sequence 990, Application US/10995561
; Publication No. US2005027054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559

```

APPLICANT: CARGILL, Michele et al.
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
TITLE OF INVENTION: CARLIOVASCLULAR DISORDERS AND DRUG RESPONSE, METHODS OF
TITLE OF INVENTION: DETECTION AND USES THEREOF
FILE REFERENCE: CU001559

```

? CURRENT APPLICATION NUMBER: 10/10/995,561
? CURRENT FILING DATE: 2004-11-24
? NUMBER OF SEO ID NOS: 85702
? SOFTWARE: FASTSEQ FOR WINDOWS VERSION 4.0
? SEO ID NO 990
? LENGTH: 2102
? TYPE: PRT
? ORGANISM: Homo sapiens
? 10-995-561-990

```

Query Match	8.1%	Score 83.5;	DB 6;	Length 2102;
Best Local Similarity	20.7%;	Pred. No. 10;		
Matches 45;	Conservative 18;	Mismatches 63;	Indels 91;	Gaps 9;

```

QY      40  KVSISLIGRAASLRRHVI-----NNTLQQLAALRLAPALPPPLPLGSEDFLSAL 94
QY      543  KYKSLPRPFRQQQEQGLKKQQQGHQWQHQQGSAPRTVPSPSPQPV-----T 611
Db
QY      95  TGS1-----LRDLTSMDTGPQNPVT--PLGLNENVPQDDPVFLS----- 135
QY      632  LGAVPAQAPAPPPPKALYPGALGRPPMPNMFDPRMM--IPPYDPLRLLQRPPLDFY 669
Db
QY      136  -----ALSRXYLDSC-----LDDFF----- 151
QY      690  PPGVHPSGLVPERSDSGSSSPPDRHAAPAMLRBGTPEVPDPKLAMVGQVFTATPAER 749
Db
QY      152  -----LDIDTSAVEKEPARAPPEPHNLFCAPG 179
Db      750  PLTSPLRQAADDDKGRSETPPVPPPPY-LASTPG 785

```

RESULT 10
US-10-995-561-989
; Sequence 989, Application US/10995561

```

/ APPLICANT: CARGILL, Michele et al..
/ TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
/ TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
/ TITLE OF INVENTION: DETECTION AND USES THEREOF
/ FILE REFERENCE: CL001559
/ CURRENT APPLICATION NUMBER: US/10/995,561
/ CURRENT FILING DATE: 2004-11-24
/ NUMBER OF SEQ ID NOS: 85702
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 989
/ LENGTH: 2108
/ TYPE: FRT
/ ORGANISM: Homo sapiens
/ US-10-995-561-989

```

Query Match	8.1%;	Score 83.5;	DB 6;	Length 2108;
Best Local Similarity	20.7%;	Pred. No. 10;		
Matches	45;	Conservative	18;	Mismatches 63;
				Indels 91;
				Gaps 9

[illegible]

RESULT 11


```

US-10-995-561-991
; Sequence 991, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 991
; LENGTH: 2157
; TYPE: prt
; ORGANISM: Homo sapiens
US-10-995-561-991

```

Query Match	8.1%;	Score 83.5;	DB 6;	Length 2157;
Best Local Similarity	20.7%;	Pred. No. 11;		
Matches 45;	Conservative 18;	Mismatches 63;	Indels 91;	Gaps 9;

QY	40	KVQNSIGRASTIRRHVL-----NHTTLOOLAALAPALPPPLFGEEDFISAT	94
Db	638	KYQNSLPRFRQKQOQEQLLKQOQKQHQWQHQGGSAPPTVPBPBPQV-----T	686
QY	95	IGSI-----LRBLDTSMDGTBPQNVT--PLGIONEVPPQDPVFL-----	135
Db	687	LGAVPAPQAPPPPRKALYGALGRPPMPNMFDPMMN--TPPVYDPRLLQGRPLDFY	744
QY	136	ALSSRYLQDSC-----LDDFF-----	151
Db	745	PCGVHPSGLVPRERSDSCGSSSBPDRHAPAMLREKGTTPVDPKLAWGVDFATPAEPR	804
QY	152	-----LDDTSAVEKEPRAPRPBPRLFCARG	179
Db	805	PLTSEPLRQADDEDDKQMSSETPVPVPPPPY--LASYDG	840

```

      RESULT 12
US-10-857-780-23
      Sequence 23, Application US/10857780
      Publication No. US20050272043A1
      GENERAL INFORMATION:
      APPLICANT: ROTH, RICHARD B.
      APPLICANT: BRAUN, ANDREAS
      APPLICANT: KAMMERER, STEFAN M.
      APPLICANT: NELSON, MATTHEW ROBERTS
      APPLICANT: RENELAND, RIKARD HENRY
      APPLICANT: HOVAL-WRIGHTSON, CAROLYN R.
      TITLE OF INVENTION: METHODS FOR IDENTIFYING RISK OF BREAST CANCER AND TREATMENTS
      TITLE OF INVENTION: THEBOP
      FILE REFERENCE: SEQ.4069-CP
      CURRENT APPLICATION NUMBER: US/10/857,780
      CURRENT FILING DATE: 2004-05-28
      PRIOR APPLICATION NUMBER: 10/723,681
      PRIOR FILING DATE: 2003-11-25
      PRIOR APPLICATION NUMBER: 60/490,234
      PRIOR FILING DATE: 2003-07-24
      PRIOR APPLICATION NUMBER: 60/525,239
      PRIOR FILING DATE: 2003-11-25
      NUMBER OF SEQ ID NOS: 4962
      SOFTWARE: PatentIn version 3.2
      SEQ ID NO 23
      LENGTH: 2101
      TYPE: FRT
      ORGANISM: Homo sapiens
US-10-857-780-23

```

Query Match	7.9%	Score 82;	DB 6;	Length 2101;
Best Local Similarity	27.1%	Pred. No. 14;		
Matches	45;	Conservative	21;	Mismatches 76;
				Indels 24;
				Gaps 7

```
QY      5 LKRGSHPLJEEBEE--WEMSPAGIQQOALRLISLKVORSIGPRAPSLRRHYLIHNTL 62
       : ::::: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db     1627 LOENKEKLRAEAEHLGHELOQAUGLKTEAOTCHILTAQVRSLEAQVAHADDOQLRDIGKF 1666
       : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      63 QQQLAAFLRLAPARLPPEPLFLGEEDFSL-SATGISILRELD-TSMGDTEPPONVPYPLG 120
       : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db     1687 QVATDAKSKBPQKPOLDSTISLDSCGEGLPILSTSKLPQPQDGTSVPGEPASP-- 1744
       : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      121 LQNEVPPQDPDVFFLEALSRYLG-----DSGLD---DFELD 153
       : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db     1745 ISQRUPRK-----VESLESLYPTPIPARSQAPLESLSIDSGDVFLD 1785
```

```

RESULT 13
US-11-134-563-2
: Sequence 2, Application US/11134563
: Publication NO. US20050287563A1
: GENERAL INFORMATION:
: APPLICANT: Leong, John M.
: APPLICANT: Campellone, Kenneth G.
: TITLE OF INVENTION: ESPV NUCLEIC ACIDS AND PROTEINS AND
: TITLE OF INVENTION: USES THEREOF
: FILE REFERENCE: 07917-280001
: CURRENT APPLICATION NUMBER: US/11/134,563
: CURRENT FILING DATE: 2005-05-20
: PRIOR APPLICATION NUMBER: US 60/573,600
: PRIOR FILING DATE: 2004-05-20
: NUMBER OF SEQ ID NOS: 26
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 2
: LENGTH: 364
: TYPE: PRT
: ORGANISM: Escherichia coli
: US-11-134-563-2

```

Query Match	7.9%	Score 81.5;	DB 7;	Length 384;
Best Local Similarity	23.5%;	Pred. No. 2;		
Matches 38;	Conservative 24;	Mismatches 51;	Indels 49;	Gaps 8;

```

QY 4 QRSIGRAASLSLRHVLIHNHTLOOLANALRLA-----PAPALF-PPPLPLGEGDEPFLSATIG 96
Db 73 ESSLHQQLNPNQRQRIQLHAEHGIKPASMMEHIIPAPNWPAPPPVQNEQSRPLPVAQ 132
QY 97 SLIRELDTSMGTETPPON-----PATPLGLONEVP-PPDDVFLEALSRLYGD 144
Db 133 RLIVQHL--AEHGIGAPAKMAHEHIIPAPNWPAPPLPVQNEQSRPLPD---VAQRVQHLAE 187
QY 145 SGLDLFFLDIDTSAVKEBPAP-----APPEPPPN 173
Db 188 HGI-----QPARSMAEHIIPAPNWPAPPPVON 215

```

```

RESULT 14
US-11-205-109-18
Sequence 18, Application US/11205109
Publication No. US20050287641A1
GENERAL INFORMATION:
APPLICANT: Parmec, Chris
APPLICANT: Zazopoulos, Emmanuel
APPLICANT: Saftia, Alfredo
TITLE OR INVENTION: GENE CLUSTER FOR RAMOPLANTIN BIOSYNTHESIS
FILE REFERENCE: 3002-205
CURRENT APPLICATION NUMBER: US/11/205,109
CURRENT FILING DATE: 2005-08-17
PRIOR APPLICATION NUMBER: US/09/976,059
PRIOR FILING DATE: 2001-10-15
PRIOR APPLICATION NUMBER: US 60/339,924
PRIOR FILING DATE: 2000-10-13
NUMBER OF SEQ ID NOS: 46
SOFTWARE: PatentIn version 3.0
SEQ ID NO 18
LENGTH: 891

```

TYPE: PRT
ORGANISM: Actinoplanes sp.
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)_(1)
OTHER INFORMATION: V is a non-standard initiator codon. It is expected that the big
OTHER INFORMATION: synthesized protein will have a formylmethionine residue at this
US-11-205-109-18

Query Match 7.8%; Score 81; DB 7; Length 891;
Best Local Similarity 27.5%; Pred. No. 6.2; Indels 22; Gaps 7;
Matches 38; Conservative 17; Mismatches 61

QY 41 VQSGPPAPSLRRHVLHNTLQOALRLAP---APALPPEP-----LFLGEBDFSLSA 93
DB 34 IERLLAAMPFMHNVAR---PQDEALRRERRLAPVPEPPARAVLAVDGSADL 90
QY 94 TIGSLRELD-TSMGTEPPONVPTPLQNEVPPQDPVFLFALSRYLGSDGLDFFL 152
DB 91 VLVARDLRDLALALAPRR--APGRKPAPPPAPPP---SAAPAWGLGCGPDRWA 145
QY 153 DITSVKEPPAPPP 170
DB 146 EL-----RVPARGPADP 157

RESULT 15

US-10-055-877-58
Sequence 58, Application US/10055877
Publication No. US20050288241A1
GENERAL INFORMATION:
APPLICANT: Decristofaro, Marc
APPLICANT: Padigaru, Muralidhara
APPLICANT: Miller, Charles
APPLICANT: Tchernev, Velizar
APPLICANT: Zhong, Mei
APPLICANT: Anderson, David
APPLICANT: Ballinger, Robert
APPLICANT: Gerlach, Valerie
APPLICANT: Spytek, Kimberly
APPLICANT: Rattelli, Luca
APPLICANT: Kekuda, Ramesh
APPLICANT: Guo, Xiaojia
APPLICANT: Zernusen, Bryan
APPLICANT: Andrew, David
APPLICANT: Mezes, Peter
APPLICANT: Patturajan, Meera
APPLICANT: Burgess, Catherine
APPLICANT: Eilsen, Andrew
APPLICANT: Wolenc, Adam
APPLICANT: Baumgartner, Jason
APPLICANT: Shinkets, Richard
APPLICANT: Gusev, Vladimir
APPLICANT: Vernet, Corine
APPLICANT: Taupier Jr., Raymond
APPLICANT: Pena, Carol
APPLICANT: Shenoy, Suresh
APPLICANT: Li, Li
APPLICANT: Caeman, Stacie
APPLICANT: Boldog, Ferenc
TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoded Thereby
FILE REFERENCE: 21402-251
CURRENT APPLICATION NUMBER: US/10/055,877
CURRENT FILING DATE: 2002-01-22
PRIOR APPLICATION NUMBER: 60/262,892
PRIOR FILING DATE: 2001-01-19
PRIOR APPLICATION NUMBER: 60/263,598
PRIOR FILING DATE: 2001-01-23
PRIOR APPLICATION NUMBER: 60/263,799
PRIOR FILING DATE: 2001-01-24
PRIOR APPLICATION NUMBER: 60/264,117
PRIOR FILING DATE: 2001-01-25

PRIOR APPLICATION NUMBER: 60/264,139
PRIOR FILING DATE: 2001-01-25
PRIOR APPLICATION NUMBER: 60/264,478
PRIOR FILING DATE: 2001-01-26
PRIOR APPLICATION NUMBER: 60/263,351
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/272,870
PRIOR FILING DATE: 2001-03-02
PRIOR APPLICATION NUMBER: 60/275,990
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 60/275,927
PRIOR FILING DATE: 2001-03-14
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 512
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 58
LENGTH: 606
TYPE: PRT
ORGANISM: Homo sapiens
US-10-055-877-58

Query Match 7.8%; Score 80.5; DB 6; Length 606;
Best Local Similarity 22.5%; Pred. No. 4.3; Indels 91; Gaps 10;
Matches 48; Conservative 13; Mismatches 61

QY 3 GGLRKHSLEEEB-----RWMSFAGISTQOALLRISLKVQSLG-- 46
DB 4 GGVHTSVPLSBOEAPQTRAGTASAGTGRAPRGAP-----PLSBOSSGCG 55
QY 47 -----PRAPSLRRHVLHNTLQOALRLAPALPPEPFLGEBDFSLATIGILR 100
DB 56 GTSRRPPPPAGKGLSLALRLRG-----PPRAMVLSQEE----- 93
QY 101 ELDTSMGTEPPONVPTPLQNEVPPQDPVFLFALSRYLGSDGLDFFLDTSAVE 160
DB 94 --PDSARTSEAQ-PLGPAFTGAAPPPGPP-----SDSBEAAVE 130
QY 161 K-----EPARA---PEPPHNTFCAPGSWEW 183
DB 131 KVEVLAQPAATAPHPPEPPS-----GGMGW 157

Search completed: February 6, 2006, 10:16:56
Job time : 18 secs